

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 19:04:17 ; Search time 4194.54 Seconds
(without alignments)
10726.543 Million cell updates/sec

Title: US-09-888-358-2
Perfect score: 1546
Sequence: 1 ggctaggcgccgcgagcg.....gtctaaagatgatgaacttc 1546

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

To number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1546	100.0	1546	6	AX403085
2	1546	100.0	1556	9	AF151827
3	1504.8	97.3	1579	9	BC001398
4	1498	96.9	1560	9	BC009330
5	1497	96.8	1550	9	HS0801436
6	1493.8	96.6	1554	9	AK026060
7	1443.8	93.4	1643	6	AR116038
8	1381.2	89.3	1402	17	AF119864
9	1054.4	68.2	1114	6	AX403084
10	1054.4	68.2	1114	9	AF317711
11	786.8	50.9	173434	2	AC098934
12	551.2	35.7	195701	2	AC124203
13	537.2	34.7	156175	2	AC111356
14	530	34.3	128802	2	AC123139
15	522	33.8	537	6	AX106606
16	522	33.8	537	6	AX140897
17	522	33.8	537	6	AX200757
18	522	33.8	537	6	AX267413
19	520	33.6	55252	2	AC119489
20	494	32.0	139488	9	AC003043
21	494	32.0	180520	2	AC019152
22	470.4	30.4	189391	2	AC019151
23	458.4	29.7	177613	2	AC123014
24	445.4	28.8	452	6	AX408486
25	437.4	28.3	758	6	AX421524
26	406.6	26.3	194025	2	AC114721
27	369	23.9	384	9	BC011178
28	296.8	19.2	138839	2	AC108295
29	242.2	15.7	1284	9	BC027322
30	240.6	15.6	1517	9	AF125531
31	218.4	14.1	240	6	AX408532
32	210.4	13.6	162456	2	AC099097
33	174.8	11.3	55774	2	AC091303
34	170.2	11.0	512	6	AX400653
35	166.6	10.8	82586	2	AC096040
36	166.6	10.8	188670	2	AC127606
37	154.4	10.0	194025	2	AC114721
38	143.8	9.3	195768	2	AC068775
39	140.4	9.1	165490	2	AC025326
40	140.4	9.1	190430	2	AC087873
41	140.4	9.1	219037	10	AL596258
42	139.6	9.0	197197	2	AC118439
43	138.8	9.0	217422	2	AC116713
44	130.4	8.4	3070	3	AY070516
45					

ALIGNMENTS

RESULT 1
AX403085
LOCUS AX403085
DEFINITION Sequence 2 from Patent WO0198355.
ACCESSION AX403085
VERSION AX403085.1 GI:21388034
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Lewin,D., Adams,S.H. and Yu,X.X.
TITLE Cgi-69 compositions and methods of use
JOURNAL Patent: WO 0198355-A 2 27-DEC-2001;

Genentech, Inc. (US) ; Curagen Corporation (US)

Location/Qualifiers

1. .1546

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 272 a 489 c 459 g 326 t

ORIGIN

Query Match 100.0%; Score 1546; DB 6; Length 1546;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTAGTGGCTGCGAGCGCGCGAGCCACGAGGCGGACGAGTAAATGGCGCCGCTT 60
 Db 1 GGCTAGTGGCTGCGAGCGCGCGAGCGCGAGCGGACGAGTAAATGGCGCCGCTT 60

QY 61 GGCCTTGGCGCGCGCGCGAGCGGACGAGGCGGACGAGTAAATGGCGCCGCTT 120
 Db 61 GGCCTTGGCGCGCGCGCGAGCGGACGAGGCGGACGAGTAAATGGCGCCGCTT 120

QY 121 GCTGACGAGGACCTGCGGCGATCAGCCCTCCAGCAAAATGGTGGCTCAGGACCGGG 180
 Db 121 GCTGACGAGGACCTGCGGCGATCAGCCCTCCAGCAAAATGGTGGCTCAGGACCGGG 180

QY 181 GCTGTGTTACCTCTCTCTTATGACACCCCTGAGCGTGGTGAAGTTCCGCTGAGTCT 240
 Db 181 GCTGTGTTACCTCTCTCTTATGACACCCCTGAGCGTGGTGAAGTTCCGCTGAGTCT 240

QY 241 CAGCGCCCTCCATGGCGAGCGAGCTGATGCTTCTCCAGCTGTGGAGCTTCTCTAT 300
 Db 241 CAGCGCCCTCCATGGCGAGCGAGCTGATGCTTCTCCAGCTGTGGAGCTTCTCTAT 300

QY 301 ACCAAATGGAAGTGCCTCTCTTATGCAATGGTGTCTGGAGCTCTGACCTGAGTCC 360
 Db 301 ACCAAATGGAAGTGCCTCTCTTATGCAATGGTGTCTGGAGCTCTGACCTGAGTCC 360

QY 361 AATGTGCGCCCTGTGCGACCTGGTTTCAAGACCTACCGCTTCACTGGCAGTATGAT 420
 Db 361 AATGTGCGCCCTGTGCGACCTGGTTTCAAGACCTACCGCTTCACTGGCAGTATGAT 420

QY 421 GCCTTCGTGAGATCGTGGCGACGAGGCGTCCAGGACCTCTGGAGCGGCTTCCCGCC 480
 Db 421 GCCTTCGTGAGATCGTGGCGACGAGGCGTCCAGGACCTCTGGAGCGGCTTCCCGCC 480

QY 481 ACCCTGTGATGATGCTGCGACCTACCGCCATCTACTTCACTGATGATGATGATGAT 540
 Db 481 ACCCTGTGATGATGCTGCGACCTACCGCCATCTACTTCACTGATGATGATGATGAT 540

QY 541 GCCTTCGTGATGCTGCGACCTACCGCCATCTACTTCACTGATGATGATGATGATGAT 600
 Db 541 GCCTTCGTGATGCTGCGACCTACCGCCATCTACTTCACTGATGATGATGATGATGAT 600

QY 601 CTGGCCGCTGCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Db 601 CTGGCCGCTGCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

QY 661 CAGGCTCAGCATGCTGACCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db 661 CAGGCTCAGCATGCTGACCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

QY 721 GGTGGCTGGGCTCACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 780
 Db 721 GGTGGCTGGGCTCACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 780

QY 781 TCAGCCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 TCAGCCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

QY 841 AAGGACAGACTTCTGTGGGCTGAGCTTTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 900
 Db 841 AAGGACAGACTTCTGTGGGCTGAGCTTTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 900

QY 901 GCAGTGTGACTCTACCTTTGAGCTGTTAAAGACCCCAAGCCAGGTCGCTCTGGGAGCG 960

Db 901 GCAGTGTGACTCTACCTTTGAGCTGTTAAAGACCCCAAGCCAGGTCGCTCTGGGAGCG 960
 QY 961 ATGAGAGCTGTGAGAGTGAACCCCTGATGTGACTTCCACCTGGCTGCTGCTGGGAGG 1020
 Db 961 ATGAGAGCTGTGAGAGTGAACCCCTGATGTGACTTCCACCTGGCTGCTGCTGGGAGG 1020

QY 1021 ATCCGCGCCGAGTCGGGACACAGGAGCTTTTTCAGAGGCTTCTTCTCGGATCATCAAG 1080
 Db 1021 ATCCGCGCCGAGTCGGGACACAGGAGCTTTTTCAGAGGCTTCTTCTCGGATCATCAAG 1080

QY 1081 GCTGCCCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 Db 1081 GCTGCCCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140

QY 1141 AGGCTGAACAGGACCGGCTTCTGGGCGCTGAAAGGGGCAAGGAGGCAAGCCCGTC 1200
 Db 1141 AGGCTGAACAGGACCGGCTTCTGGGCGCTGAAAGGGGCAAGGAGGCAAGCCCGTC 1200

QY 1201 TCTCCACAGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
 Db 1201 TCTCCACAGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260

QY 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
 Db 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320

QY 1321 CTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
 Db 1321 CTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380

QY 1381 ATCACTTACCAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
 Db 1381 ATCACTTACCAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

QY 1441 CTGTGTTGCTGTGAGTGGGAGTGTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
 Db 1441 CTGTGTTGCTGTGAGTGGGAGTGTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500

QY 1501 CTCCCTGAGGCTGTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAA 1546
 Db 1501 CTCCCTGAGGCTGTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAA 1546

AF151827 1556 bp mRNA linear PRI 18-MAY-2000
 Homo sapiens CGI-69 protein mRNA, complete cds.

AF151827.1 GI:4929606

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

118. .1173

CDS

Query Match	100.0%;	Score 1546;	DB 9;	Length 1556;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 1546;	Conservative	0;	Mismatches	0; Indels
				0; Gaps
QY	1	GGCTAGGTGGCGCTGCGAGGCGCGGAGCCACGAGGGCGGACGGACGTAAATGGGCGCGGCT	60	
Db	1	GGCTAGGTGGCGCTGCGAGGCGCGGAGCCACGAGGGCGGACGGACGTAAATGGGCGCGGCT	60	
QY	61	GGCCCTGGGCGCGCGCGCGGACGAGCACGAGCTTAGAGCGAGACTGAGCTTCAGATG	120	
Db	61	GGCCCTGGGCGCGCGCGCGGACGAGCACGAGCTTAGAGCTTAGAGCTTCAGATG	120	
QY	121	GCTGACAGGACCCCTCGGGCATCAGGCCCTCAGCAAAATGGTGGCTCAGGACACGGG	180	
Db	121	GCTGACAGGACCCCTCGGGCATCAGGCCCTCAGCAAAATGGTGGCTCAGGACACGGG	180	
QY	181	GCTGTGGTACTCTCTCTTCATGACACCCCTGAGCGTGGTGAAGGTCGCTGAGCT	240	
Db	181	GCTGTGGTACTCTCTCTTCATGACACCCCTGAGCGTGGTGAAGGTCGCTGAGCT	240	
QY	241	CAGCGGCCCTCCATGCGCCAGCGAGTGATGCTTCCTCCAGACTGTGGAGCTCTCCTAT	300	
Db	241	CAGCGGCCCTCCATGCGCCAGCGAGTGATGCTTCCTCCAGACTGTGGAGCTCTCCTAT	300	
QY	301	ACCAAATGGAAGTGCCCTGCTGATTGCAATGGTGTCTCTGGAGCCTCTGACTGTCGCCA	360	
Db	301	ACCAAATGGAAGTGCCCTGCTGATTGCAATGGTGTCTCTGGAGCCTCTGACTGTCGCCA	360	
QY	361	AATGTGCCGCTGTGCCACCTGTTTCAAGACCCCTACCCGCTTCACGTGGACCATGGAT	420	
Db	361	AATGTGCCGCTGTGCCACCTGTTTCAAGACCCCTACCCGCTTCACGTGGACCATGGAT	420	
QY	421	GCCTTCGTGAAGATCGTGTAGGCGACGAGGGCACGAGACCCCTCTGGAGCGGCTCCCGGC	480	
L	421	GCCTTCGTGAAGATCGTGTAGGCGACGAGGGCACGAGACCCCTCTGGAGCGGCTCCCGGC	480	
QY	481	ACCTGTGTGATGACTGTGCCAGCTACGCGCATCTTCACTGCGCTATGACCAACTGAAG	540	
Db	481	ACCTGTGTGATGACTGTGCCAGCTACGCGCATCTTCACTGCGCTATGACCAACTGAAG	540	
QY	541	GCCTTCCTGTGTGGTGTGAGCCCTGACCTCTGACCTTAGCGACCCATGGTGGTGGCGG	600	
Db	541	GCCTTCCTGTGTGGTGTGAGCCCTGACCTCTGACCTTAGCGACCCATGGTGGTGGCGG	600	
QY	601	CTGGCCCGCTGGGCAACGTGTGATGATGAGCCCTGTGGACTTATGCGGACAAAGCTG	660	
Db	601	CTGGCCCGCTGGGCAACGTGTGATGATGAGCCCTGTGGACTTATGCGGACAAAGCTG	660	
QY	661	CAGGCTCAGCATGTGTGCTACCGGAGCTGGGTGCCTGTGTTCGAACCTGCACTGGGCTCAG	720	
Db	661	CAGGCTCAGCATGTGTGCTACCGGAGCTGGGTGCCTGTGTTCGAACCTGCACTGGGCTCAG	720	
QY	721	GGTGGCTGGGCTCACTGTGGCTGGCTGGGGCCCACTGCCCTTCGAGATGTGCCCTTC	780	
Db	721	GGTGGCTGGGCTCACTGTGGCTGGCTGGGGCCCACTGCCCTTCGAGATGTGCCCTTC	780	
QY	781	TCAGCCCTGACTGGTTCAACTATAGAGCTGGTGAAGAGCTGGCTCAATGGGTCAGCGCG	840	
Db	781	TCAGCCCTGACTGGTTCAACTATAGAGCTGGTGAAGAGCTGGCTCAATGGGTCAGCGCG	840	


```

Db 1450 TTCCCTGTTGTTGCTAGCTGGGCAATGCTCCAGGAACCAAGAGCCCTCAGCCTGGTG 1509
QY 1497 TAGTCTCCCTGACCCCTGTTAAATTCCTTAAAGTCTAAAGATGATGAATCTTC 1546
Db 1510 TAGTCTCCCTGACCCCTGTTAAATTCCTTAAAGTCTAAAGATGATGAATCTTC 1559

RESULT 4
BC009330
LOCUS BC009330
DEFINITION Homo sapiens, clone MGC:16684 IMAGE:4126268, mRNA, complete cds.
ACCESSION BC009330
VERSION BC009330.1 GI:14424605
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1560)
Strausberg,R.
Direct Submission
Submitted (12-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dierich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 26 Row: g Column: 12.
Location/Qualifiers
1. .1560
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:16684 IMAGE:4126268"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC 19"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
103..1158
/codon_start=1
/product="Unknown (protein for MGC:16684)"
/protein_id="AAH09330.1"
/db_xref="GI:14424606"
/translation="MAQDFAGISPLQQWVAGTGVVTSIFMTPLDVKVRLQSQRP
SVASEMFSSRLWSISYKWKCLLYCNGVLELYLCPNGARCATWFODFRFTGMDA
FVKLVHSGTTLTWSGLPATLVMTVPATAIYFTADQLKAFICGRALTSDLVAPWVAG
ALARGTWTVI.SPLELMRTKLOQHVSYRELCAVCRTAVAGGWSRLWLGFTALRD
VPFSALYWFNVELVKSMLNGLRPKDQTSVGMSPFVAGISGVAAVUTLFFDVFVQTRQ
VALGAEAVRYNPLVDSTWLLRLRRFRAESGKGLFAGFLPRIKAAPSCAEMISTYE
FKGSFQRLNDRLIG"
296 a 488 g 323 t
BASE COUNT
ORIGIN
Query Match
96.9%; Score 1498; DB 9; Length 1560;

```

```

Best Local Similarity 99.3%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 33 AGGCGGCGAGCGAAGCTTAATGGGCGCGCTGGCCCTGGGCGCGCGCGAGGACGACG 92
Db 18 AGGCGGCGGAGCGGCTATGGGCGCGCTGGCCCTGGGCGCGCGCGAGGACGACG 77

QY 93 CTAGAGCCAGGACTGAAGCTTCAAGATGGCTGAACAGGACCCCTGCGGGCATCAGCC 152
Db 78 CTAGAGCCAGGACTGAAGCTTCAAGATGGCTGAACAGGACCCCTGCGGGCATCAGCC 137

QY 153 CCAGCAAAATGCTGCTCAGGACACCGGCGCTGTGTTACCTCTCTTCTATGACACCC 212
Db 138 CCAGCAAAATGCTGCTCAGGACACCGGCGCTGTGTTACCTCTCTTCTATGACACCC 197

QY 213 GGAGCTGGTGAAGCTTGGCTCTCAGCTCTCAGCGGCGCTCCATGCGCGAGGACGATGCC 272
Db 198 GGAGCTGGTGAAGCTTGGCTCTCAGCTCTCAGCGGCGCTCCATGCGCGAGGACGATGCC 257

QY 273 TTCCTCCAGACTGTGGAGCTCTCCTATACCAATGGAAGTGCCTCCTGTATTGCAATGG 332
Db 258 TTCCTCCAGACTGTGGAGCTCTCCTATACCAATGGAAGTGCCTCCTGTATTGCAATGG 317

QY 333 TGTCTGTGAGCCCTCTGTACCTGTGCGCCAAATGCTGCCCGCTGTGCCACTGTGTTTCA 392
Db 318 TGTCTGTGAGCCCTCTGTACCTGTGCGCCAAATGCTGCCCGCTGTGCCACTGTGTTTCA 377

QY 393 CCCTACCCGCTTCACTGGCACCATGGATGCTTCTGTAAGATCGTGAGGACGAGGCGAC 452
Db 378 CCCTACCCGCTTCACTGGCACCATGGATGCTTCTGTAAGATCGTGAGGACGAGGCGAC 437

QY 453 CAGACCCCTGTGAGGCGCTCCCGGCAACCTGTGATGACTGTGCCAGTACCGGCAT 512
Db 438 CAGACCCCTGTGAGGCGCTCCCGGCAACCTGTGATGACTGTGCCAGTACCGGCAT 497

QY 513 CTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTCAGCCCTGACCTCTGA 572
Db 498 CTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTCAGCCCTGACCTCTGA 557

QY 573 CCTCTAGCACCATGTGTGGCTGGCGCGCTGGCGCGCGCTGGGCAACCGTGTGATCAG 632
Db 558 CCTCTAGCACCATGTGTGGCTGGCGCGCTGGCGCGCGCTGGGCAACCGTGTGATCAG 617

QY 633 CCCCCTGGAGCTTATGCGGCAAAAGTTCAGGCTCAGCATGTGTCTACCGGAGCTGGG 692
Db 618 CCCCCTGGAGCTTATGCGGCAAAAGTTCAGGCTCAGCATGTGTCTACCGGAGCTGGG 677

QY 693 TGCTGTGTTCGAACTGCTGCTCAGGCTGGCTGGCGCTCAGTGTGGCTGGGCTGGG 752
Db 678 TGCTGTGTTCGAACTGCTGCTCAGGCTGGCTGGCGCTCAGTGTGGCTGGGCTGGG 737

QY 753 CCCCCTGCGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATCAGCTGT 812
Db 738 CCCCCTGCGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATCAGCTGT 797

QY 813 GAAGAGCTGGCTCAATGGTTCAGGCCGGAAGCAACAGACTTCTGTGGGCATGAGTTGT 872
Db 798 GAAGAGCTGGCTCAATGGTTCAGGCCGGAAGCAACAGACTTCTGTGGGCATGAGTTGT 857

QY 873 GGCTGTGGCATCTCAGGAGCGGTGGCTGCACTGTGCTACCTTTGACGTGTTAAA 932
Db 858 GGCTGTGGCATCTCAGGAGCGGTGGCTGCACTGTGCTACCTTTGACGTGTTAAA 917

QY 933 GACCAACGCGCAGTGTGCTCTGGGAGCGCATGAGGCTGTGAGAGTCAACCCCTCATGT 992
Db 918 GACCAACGCGCAGTGTGCTCTGGGAGCGCATGAGGCTGTGAGAGTCAACCCCTCATGT 977

QY 993 GGACTCCACCTGGCTGCTGCTGGGAGGATCCGGGCCGAGTCCGGGCAACCAAGGACTCTT 1052
Db 978 GGACTCCACCTGGCTGCTGCTGGGAGGATCCGGGCCGAGTCCGGGCAACCAAGGACTCTT 1037

QY 1053 TGCAGGCTTCTCTTCCTCGCATCATCAAGGCTGCCCCCTCTCTGTGCCATCATGATCAGCAC 1112

```


Db 554 GTGAGCGCTGACCTCTGACCTTACGACCCATGTGTGCTGGCGCGTGGCCCGCTGG 613
 QY 614 GCACCTGACTGTGATGACGCCCTGTGAGCTTATGCGGACAAAGCTGAGGCTCAGCATG 673
 Db 614 GCACCTGACTGTGATGACGCCCTGTGAGCTTATGCGGACAAAGCTGAGGCTCAGCATG 673
 QY 674 TGTCTACCGGAGCTGGTGTGCTGCTTGAACCTGAGTGTGCTCAGGCTGGTGGCT 733
 Db 674 TGTCTACCGGAGCTGGTGTGCTGCTTGAACCTGAGTGTGCTCAGGCTGGTGGCT 733
 QY 734 CACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 793
 Db 734 CACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 793
 QY 794 GGTTCACCTATGAGCTGTGAGAGCTGGCTCAATGGGTTCAGGCCGAGGACGACAT 853
 Db 794 GGTTCACCTATGAGCTGTGAGAGCTGGCTCAATGGGTTCAGGCCGAGGACGACAT 853
 QY 854 CTGTGGGCTAGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 Db 354 CTGTGGGCTAGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 QY 914 TACCTTTGACCTGTGTAAGACCCACCGCCAGCTGCTCTGGAGCGATGAGGCTGTGA 973
 Db 914 TACCTTTGACCTGTGTAAGACCCACCGCCAGCTGCTCTGGAGCGATGAGGCTGTGA 973
 QY 974 GAGTGAACCCCTGATGTGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1033
 Db 974 GAGTGAACCCCTGATGTGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1033
 QY 1034 CGGGACCAAGGAGCTTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
 Db 1034 CGGGACCAAGGAGCTTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
 QY 1094 GTGCCATCATGATCAGCACTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGACCCAGG 1153
 Db 1094 GTGCCATCATGATCAGCACTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGACCCAGG 1153
 QY 1154 ACCGCTTCTGGGCGCTGAAAGGGGCAAGGAGGCAAGGACCCGCTCTTCCACGGATG 1213
 Db 1154 ACCGCTTCTGGGCGCTGAAAGGGGCAAGGAGGCAAGGACCCGCTCTTCCACGGATG 1213
 QY 1214 GGGAGGGGAGGAGGAGCCAGCCAGTGTCTTCTCAGAGCTGAGGAGGAGGCT 1273
 Db 1214 GGGAGGGGAGGAGGAGCCAGCCAGTGTCTTCTCAGAGCTGAGGAGGAGGCT 1273
 QY 1274 TGTTCCTTCTTCCCTCCGCGCAAGACTCCAGGCGAGGCTGTCCCTTGGCGGCCAG 1333
 Db 1274 TGTTCCTTCTTCCCTCCGCGCAAGACTCCAGGCGAGGCTGTCCCTTGGCGGCCAG 1333
 QY 1334 CACTTCTCAGACACACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1393
 Db 1334 CACTTCTCAGACACACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1393
 QY 1394 CCCCCAAGTTCAGAGCAAACTTCCAGCTGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1453
 Db 1394 CCCCCAAGTTCAGAGCAAACTTCCAGCTGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1453
 QY 1454 AGCTGGGATGTCTCAGGAACCAAGAGCCCTCAGGCTGGTGTGCTGCTGCTGCTGCTGCT 1513
 Db 1454 AGCTGGGATGTCTCAGGAACCAAGAGCCCTCAGGCTGGTGTGCTGCTGCTGCTGCTGCT 1513
 QY 1514 GTTAATTCCTTAAGTCTAAGATGATGAA 1542
 Db 1514 GTTAATTCCTTAAGTCTAAGATGATGAA 1542
 RESULT 6
 AK026060
 LOCUS 1554 bp mRNA linear PRI 29-SEP-2000
 DEFINITION Homo sapiens cDNA: FLJ22407 fis, clone HRC08407.
 AK026060
 ACCESSION
 VERSION AK026060.1 GI:10438782

KEYWORDS

oligo capping: fis (full insert sequence).
 Homo sapiens primary human renal epithelial cells cDNA to mRNA,
 clone lib:HRC clone:HRC08407.

SOURCE

Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

ORGANISM

1 (sites)
 Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
 Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
 Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished

REFERENCE

2 (bases 1 to 1554)
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639,
 Japan [E-mail:cdna@ims.u-tokyo.ac.jp. Tel:81-3-5449-5286,
 Fax:81-3-5449-5416]

AUTHORS

1 (sites)
 Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
 Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
 Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished

TITLE

Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639,
 Japan [E-mail:cdna@ims.u-tokyo.ac.jp. Tel:81-3-5449-5286,
 Fax:81-3-5449-5416]

JOURNAL

Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639,
 Japan [E-mail:cdna@ims.u-tokyo.ac.jp. Tel:81-3-5449-5286,
 Fax:81-3-5449-5416]

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction; 5'- & 3'-end one pass sequencing; Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES

Location/Qualifiers
 1..1554
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HRC08407"
 /cell_type="primary human renal epithelial cells"
 /clone_lib="HRC"
 /notes="Cloning vector pME18SFL3"
 118..1173
 /notes="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15341.1"
 /db_xref="GI:10438783"

CDS

/translation="MADDDPAGISPLQMVASGTGAVTSLFWTLDDVKKVLSQSRP
 SMASLMBSSRLWSISYTKWCLLYCNVLEPLFCNPARGCATWFOPTFTGTMDA
 FVKIVHEGTRTLNGDPLATVMTVPATAYFTAYDQLKAPLCGRATSLDLYAPMVAG
 ALARLGTVTISPLMETKLIQAQHSYRELGAQVTRVAQGWESLWGLGPTALRD
 VPSALYWNFYELVKSMLNGPRKDOTSVGMSFVAGISGTVAALVTLPDVKTORQ
 VALGMEAVRVNPLHVDSTWLLRRIRAESGTLKGLFAGFLPRLIIKAPSCAIIIMTYE
 FGKSPFORLUNQDRLGG"

BASE COUNT

285 a 489 c 457 g 323 t

ORIGIN

Query Match 96.6%; Score 1493.8; DB 9; Length 1554;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1507; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY

14 GCGAGCGCGGAGCACAGGCGGACGAGCTTAATGGGCGCCCTGCGCTGGGCGCC 73

Db

14 GCGAGCGCGGAGCACAGGCGGACGAGCTTAATGGGCGCCCTGCGCTGGGCGCC 73

QY

74 GCGCGCGACGAGCACAGGCGGACGAGCTTAATGGGCGCCCTGCGCTGGGCGCC 133

Db

74 GCGCGCGACGAGCACAGGCGGACGAGCTTAATGGGCGCCCTGCGCTGGGCGCC 133

QY

134 CTGCGGGCATCAGCCCTCCAGCAAAATGTTGGCTCAGGACCGGGGCTGTGGTTACCT 193

Db

134 CTGCGGGCATCAGCCCTCCAGCAAAATGTTGGCTCAGGACCGGGGCTGTGGTTACCT 193

QY

194 CTCTCTTATGACACCCCTGGAGCTGTGTAAGTTCCCTGCGAGTCTCAGCGGCCCTCCA 253

Db

194 CTCTCTTATGACACCCCTGGAGCTGTGTAAGTTCCCTGCGAGTCTCAGCGGCCCTCCA 253

QY

254 TGGCCAGCGAGCTGTATGCTTCTCCAGACTGTGGAGCTCTCTCTATACCAATGGAAT 313

Db

254 TGGCCAGCGAGCTGTATGCTTCTCCAGACTGTGGAGCTCTCTCTATACCAATGGAAT 313

Db 254 TGGCCAGGAGCTGATGCTTCTCCAGACTGTGGAGCCTCTCTATACCAATGGAAGT 313
 QY 314 GCCTCCTGATTTGCAATGGTGTCTTGGAGCCTCTGTACCTGTGCCCCGCT 373
 Db 314 GCCTCCTGATTTGCAATGGTGTCTTGGAGCCTCTGTACCTGTGCCCCGCT 373
 QY 374 GTGCCACCTGTTTCAAGACCTTACCGCTTCACTGGGACCATGGATGCTTCTGGAAGA 433
 Db 374 GTGCCACCTGTTTCAAGACCTTACCGCTTCACTGGGACCATGGATGCTTCTGGAAGA 433
 QY 434 TCGTGAGGACGAGGAGGACAGAGACCTCTCTGGAGCGGCTTCCCGCCACCTGTGTGTA 493
 Db 434 TCGTGAGGACGAGGAGGACAGAGACCTCTCTGGAGCGGCTTCCCGCCACCTGTGTGTA 493
 QY 494 CTGTGCCAGCTACCGGCACTTACTTCACTGCTCTATGACCACTGAGGCGCTTCTGTGTG 553
 Db 494 CTGTGCCAGCTACCGGCACTTACTTCACTGCTCTATGACCACTGAGGCGCTTCTGTGTG 553
 QY 554 GTCCAGCCTGACCTCTGACCTCTTACGACCCATGGTGGTGGCGCTGGCCGCTGG 613
 Db 554 GTCCAGCCTGACCTCTGACCTCTTACGACCCATGGTGGTGGCGCTGGCCGCTGG 613
 QY 614 GCACCTGACTGTGATCAAGCCCTTGGAGCTTATGCGGACAAAGCTGACGCTCAGCATG 673
 Db 614 GCACCTGACTGTGATCAAGCCCTTGGAGCTTATGCGGACAAAGCTGACGCTCAGCATG 673
 QY 674 TGTCTGACCGGAGCTGGGTGCTGTGTGTTTGAAGTCTGAGGCTGAGGCTGGCGCT 733
 Db 674 TGTCTGACCGGAGCTGGGTGCTGTGTGTTTGAAGTCTGAGGCTGAGGCTGGCGCT 733
 QY 734 CACTGTGGCTGGGCTGGGCGCCACTCCCTTCAGATGTCCTTCTCAGCCCTGTACT 793
 Db 734 CACTGTGGCTGGGCTGGGCGCCACTCCCTTCAGATGTCCTTCTCAGCCCTGTACT 793
 QY 794 GGTTCACCTATGAGCTGTGAGAGCTGGCTCAATGGTTCAGGCGGAGGACCACTT 853
 Db 794 GGTTCACCTATGAGCTGTGAGAGCTGGCTCAATGGTTCAGGCGGAGGACCACTT 853
 QY 854 CTCTGGGCTGAGCTTGTGGCTGGTGGCTCTCAGGAGCGTGGCTGCACTGCTACTC 913
 Db 854 CTCTGGGCTGAGCTTGTGGCTGGTGGCTCTCAGGAGCGTGGCTGCACTGCTACTC 913
 QY 914 TACCCTTTGAGCTGTAAAGACCAAGCCAGCTGCTCTGGGAGCGATGGAGCTGTGA 973
 Db 914 TACCCTTTGAGCTGTAAAGACCAAGCCAGCTGCTCTGGGAGCGATGGAGCTGTGA 973
 QY 974 GAGTCAACCCCTGATGAGCTTCCAGCTGGCTGCTCTGGGAGCGATCCGGGCGAGT 1033
 Db 974 GAGTCAACCCCTGATGAGCTTCCAGCTGGCTGCTCTGGGAGCGATCCGGGCGAGT 1033
 QY 1034 CGGGACCAAGGAGCTTTTTCAGAGCTTCTTCTCGGATCATCAAGGCTGCGCCCTCT 1093
 Db 1034 CGGGACCAAGGAGCTTTTTCAGAGCTTCTTCTCGGATCATCAAGGCTGCGCCCTCT 1093
 QY 1094 GTGCCATCATGATCAGCACCTATGAGTTCGGGAAAGCTTCTTCAGAGCTGAACGAG 1153
 Db 1094 GTGCCATCATGATCAGCACCTATGAGTTCGGGAAAGCTTCTTCAGAGCTGAACGAG 1153
 QY 1154 ACCGCTTCTGGGCGGTGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1213
 Db 1154 ACCGCTTCTGGGCGGTGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1213
 QY 1214 GGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1273
 Db 1214 GGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1273
 QY 1274 TGTTCCTTCTCCCTCCCGGCGACAGCTTCCAGGAGGAGGAGGAGGAGGAGGAGG 1333
 Db 1274 TGTTCCTTCTCCCTCCCGGCGACAGCTTCCAGGAGGAGGAGGAGGAGGAGGAGG 1333
 QY 1334 CACTTCTCAGACACACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1393
 Db 1334 CACTTCTCAGACACACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1393

QY 1394 CCCCAGATTCAAGACCAATCTTCCAGCTGCCCTTCCCTGCTGCTGCTGCTGCTGCT 1453
 Db 1394 CCCCAGATTCAAGACCAATCTTCCAGCTGCCCTTCCCTGCTGCTGCTGCTGCTGCT 1453
 QY 1454 AGCTGGGATGCTTCCAGGAAACCAAGAGCCTCAGCCTGCTGCTGCTGCTGCTGCT 1513
 Db 1454 AGCTGGGATGCTTCCAGGAAACCAAGAGCCTCAGCCTGCTGCTGCTGCTGCTGCT 1513
 QY 1514 GTTATTTCTTAAGTCTTAAGATGATGAA 1542
 Db 1514 GTTATTTCTTAAGTCTTAAGATGATGAAA 1542

RESULT 7
 AR116038
 LOCUS AR116038 1643 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 68 from patent US 6132973.
 ACCESSION AR116038
 VERSION AR116038.1 GI:14096360
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1643)
 AUTHORS Lal, P., Hillman, J.L., Bandman, O., Shah, P., Au-Young, J., Yue, H.,
 Guegler, K.J. and Corley, N.C.
 TITLE Human regulatory molecules
 JOURNAL Patent: US 6132973-A 68 17-OCT-2000;
 FEATURES Location/Qualifiers
 source 1..1643
 BASE COUNT 307 a 501 c 476 g 359 t
 ORIGIN

Query Match 93.4%; Score 1443.8; DB 6; Length 1643;
 Best Local Similarity 99.9%; Pred. No. 2.8e-308;
 Matches 1445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 100 CAGGACTGAAGCTTCAAGATGGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 159
 Db 187 CAGGACTGAAGCTTCAAG: GGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 246
 QY 160 ATGGTGGCTCAGGACCGGGGCTGTGGTACCTCTCTTCATGACACACCTGGAGCTG 219
 Db 247 ATGGTGGCTCAGGACCGGGGCTGTGGTACCTCTCTTCATGACACACCTGGAGCTG 306
 QY 220 GTCAAGGTTGGCTGAGCTCTCAGCGGCGCTCCATGGCCAGGAGCTGATGCTCTCTCC 279
 Db 307 GTCAAGGTTGGCTGAGCTCTCAGCGGCGCTCCATGGCCAGGAGCTGATGCTCTCTCC 366
 QY 280 AGACTGTGAGCTCTCTCTTACCAAAATGGAAGTGGCTCTCTGATTCGAATGGTGTCTG 339
 Db 367 AGACTGTGAGCTCTCTCTTACCAAAATGGAAGTGGCTCTCTGATTCGAATGGTGTCTG 426
 QY 340 GAGCCTCTGATCTGCTCCCAAAATGGTGGCGCTGCGACCTGTGGTTCGAAGACCTTACC 399
 Db 427 GAGCCTCTGATCTGCTCCCAAAATGGTGGCGCTGCGACCTGTGGTTCGAAGACCTTACC 486
 QY 400 CGCTTCACTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 459
 Db 487 CGCTTCACTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 546
 QY 460 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 519
 Db 547 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 606
 QY 520 ACTGCTTATGACCAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 579
 Db 607 ACTGCTTATGACCAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 666
 QY 580 GCACCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 639

Db 667 GCACCCATGTGTGGCTGGCCGCTGGCCCGCTTGGGCGACCGGTGATGATCAGCCCCCTG 726
QY 640 GAGCTTATGCGCAAAAGCTGAGGCTCAGCATCTGTCTGACCGGAGCTGGGTGCTGT 699
Db 727 GAGCTTATGCGCAAAAGCTGAGGCTCAGCATCTGTCTGACCGGAGCTGGGTGCTGT 786
QY 700 GTTCGAACTGCACTGAGTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 759
Db 787 GTTCGAACTGCACTGAGTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 846
QY 760 GCGCTTCAGATGTGCTTCTCAGCCCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 819
Db 847 GCGCTTCAGATGTGCTTCTCAGCCCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
QY 820 TGGCTCAATGGTTCAGGCGCAAGACCAAGACTTCTGTGGCATGAGCTTGTGGCTGT 879
Db 907 TGGCTCAATGGTTCAGGCGCAAGACCAAGACTTCTGTGGCATGAGCTTGTGGCTGT 966
QY 880 GGCACTCTCAGGAGCGTGGCTGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
Db 967 GGCACTCTCAGGAGCGTGGCTGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
QY 940 CGCAGGTGCTCTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCACTGTGGACTCC 999
Db 1027 CGCAGGTGCTCTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCACTGTGGACTCC 1086
QY 1000 ACTGCTGCTCTCGGAGGATCCGGGCGAGTGGGCGACCAAGGACTCTTTGCGAGC 1059
Db 1087 ACTGCTGCTCTCGGAGGATCCGGGCGAGTGGGCGACCAAGGACTCTTTGCGAGC 1146
QY 1060 TTCTTCTTCGATCATCAAGGCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119
Db 1147 TTCTTCTTCGATCATCAAGGCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
QY 1120 TTCCGCAAAAGCTTCTTCCAGAGCTGAACAGGACCGGCTTCTGGCGGCTGAAGAGG 1179
Db 1207 TTCCGCAAAAGCTTCTTCCAGAGCTGAACAGGACCGGCTTCTGGCGGCTGAAGAGG 1266
QY 1180 CAAGAGGCAAGGACCCCGTCTTCCACCGATGGGAGAGGCGAGGAGACCCAGCC 1239
Db 1267 CAAGAGGCAAGGACCCCGTCTTCCACCGATGGGAGAGGCGAGGAGACCCAGCC 1326
QY 1240 AAGTGCCTTTCTCAGCATGAGGAGGAGGCTTGTTCCTTCCCTCCCGGCAAG 1299
Db 1327 AAGTGCCTTTCTCAGCATGAGGAGGAGGCTTGTTCCTTCCCTCCCGGCAAG 1386
QY 1300 TTCCAGGCGAGGCTGCTTCCCTCTGGCGGCGCCAGCATTTCTCAGACACAACTTCTCT 1359
D 1387 TTCCAGGCGAGGCTGCTTCCCTCTGGCGGCGCCAGCATTTCTCAGACACAACTTCTCT 1446
QY 1360 GCTGCTCAGTCTGGGATCATCACTTACCCACCCCGGAGTTCAAGACCAATCTTCC 1419
Db 1447 GCTGCTCAGTCTGGGATCATCACTTACCCACCCCGGAGTTCAAGACCAATCTTCC 1506
QY 1420 AGCTGCCCTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1479
Db 1507 AGCTGCCCTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
QY 1480 AAGCCCTCAGCTGGTGTAGTCTTCCCTGACCCCTGTTAACTTCTTAAGCTTAAGCAT 1539
Db 1567 AAGCCCTCAGCTGGTGTAGTCTTCCCTGACCCCTGTTAACTTCTTAAGCTTAAGCAT 1626
QY 1540 GAACCTTC 1546
Db 1627 GAACCTTC 1633

RESULT 8

AF119864

ID AF119864 standard; RNA; HUM; 1402 BP.

AC AF119864;

XX

SV AF119864.1
XX 12-MAY-2000 (Rel. 63, Created)
DT 09-MAY-2001 (Rel. 67, Last updated, Version 2)
XX Homo sapiens PRO2163 mRNA, complete cds.
XX
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
[1]
RN 1-1402
RP Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RA "Functional prediction of the coding sequences of 79 new genes deduced by
RT analysis of cDNA clones from human fetal liver";
RL Unpublished.
XX
[2]
RN 1-1402
RP Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RA Submitted (13-JAN-1999) to the EMBL/GenBank/DBJ databases.
RL Department of Experimental Hematology, Institute of Radiation Medicine,
RL Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
XX SWISS-PROT; Q9BZJ4; CG69_HUMAN.
XX
Key Location/Qualifiers
source 1..1402
/db_xref="taxon:9606"
/organism="Homo sapiens"
/clone="PLB8041"
/tissue_type="liver"
/dev_stage="Fetus"
4..1017
/codon_start=1
/db_xref="SWISS-PROT:Q9BZJ4"
/note="Predicted protein of HQ2163"
/product="PRO2163"
/protein_id="AAP6961.1"
/translation="MWASGTGAVTSLFMTPLDVVKVRLQSORPSMASELPSSRLWSL
SYTKWKLJYCMGVLEPLYLCPNGARCATWFQDPTFTGTMDFVKI VRHRTTTLMSG
LPATLVMTVPATAIYFTAYDOLKAPLGRALTSDLYAPNVAGALALCTVTVISPLELM
RTKLOAHVSVYRELGACVRTAVAGGWSLWLGWPTALRDVPPSALYFNFYEI VKSWL
NGLRPKDOTSVGMSFVAGGI SRTVAAYLTLPFDVVKIQRQVAGLAWAVRNPJLHVDST
WLLLRIRAESGTGLFAGFLPFIIRIKAAPSACIMISTYEFKSPFQRLNDRLLGG"

Sequence 1402 BP; 261 A; 432 C; 401 G; 308 T; 0 other;

Query Match 89.3%; Score 1381.2; DB 17; Length 1402;
Best Local Similarity 99.8%; Pred. No. 1.9e-294;
Matches 1383; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 157 CAAATGGTGGCTCAGCACCGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGGAC 216
Db 1 CAAATGGTGGCTCAGCACCGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGGAC 60
QY 217 GTGGTGAAGGTTGCGCTGAGTCTCAGCGGCCCTCCATGGCCAGAGCTGATCTTCC 276
Db 61 GTGGTGAAGGTTGCGCTGAGTCTCAGCGGCCCTCCATGGCCAGAGCTGATCTTCC 120
QY 277 TCCAGACTGTGGAGCTCTCTTATACCAATGGAGTGCCTCTGTTGCAATGGTCTC 336
Db 121 TCCAGACTGTGGAGCTCTCTTATACCAATGGAGTGCCTCTGTTGCAATGGTCTC 180
QY 337 CTGAGGCTCTGTACTGTGCCAAATGGTGCCTGTCGCCACCTGGTTTCAAGACCT 396
Db 181 CTGAGGCTCTGTACTGTGCCAAATGGTGCCTGTCGCCACCTGGTTTCAAGACCT 240

QY 397 ACCCGCTTCACTGGCAACCATGATGCTTGTGAGATGCTGAGGACGAGGCGACACAG 456
 Db 241 ACCCGCTTCACTGGCAACCATGATGCTTGTGAGATGCTGAGGACGAGGCGACACAG 300
 QY 457 ACCCTGTGGAGGGGCTCCCGCCACCATGCTGATGACTGTGCCAGCTACCGCCATCTAC 516
 Db 301 ACCCTGTGGAGGGGCTCCCGCCACCATGCTGATGACTGTGCCAGCTACCGCCATCTAC 360
 QY 517 TTCACTGCTATGACCACTGAGGCGCTTCTGTGTGCTGAGGCGCTGACCTGACCTC 576
 Db 361 TTCACTGCTATGACCACTGAGGCGCTTCTGTGTGCTGAGGCGCTGACCTGACCTC 420
 QY 577 TAGGCAACCATGCTGCTGGCGGCTGCGCCGCTGGGCAACGCTGATGATGATGATGATG 636
 Db 421 TAGGCAACCATGCTGCTGGCGGCTGCGCCGCTGGGCAACGCTGATGATGATGATGATG 480
 QY 637 CTGAGCTTATCGGCAAAAGCTGACGCTGACATGCTGTGATGACGAGCTGGGTGCC 696
 Db 481 CTGAGCTTATCGGCAAAAGCTGACGCTGACATGCTGTGATGACGAGCTGGGTGCC 540
 QY 697 TGTGTTGCAACTGACGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 756
 Db 541 TGTGTTGCAACTGACGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 757 ACTGCTTTCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 816
 Db 601 ACTGCTTTCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 817 AGCTGGCTCAATGGGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 876
 Db 661 AGCTGGCTCAATGGGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 720
 QY 877 GGTGGCTTCAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
 Db 721 GGTGGCTTCAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 QY 937 CAACGCCAGTTCGCTTGGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 996
 Db 781 CAACGCCAGTTCGCTTGGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 840
 QY 997 TCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056
 Db 841 TCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 QY 1057 GCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116
 Db 901 GCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 QY 1117 GAGTTGGCAAAAGCTTCTTCCAGAGGCTGAAACGAGGCTGAAACGAGGCTGAAACG 1176
 Db 961 GAGTTGGCAAAAGCTTCTTCCAGAGGCTGAAACGAGGCTGAAACGAGGCTGAAACG 1020
 QY 1177 GGGCAAGGAGCAAGGAGGCTGCTTCCAGGATGGGAGGAGGAGGAGGAGGAGGAGGAG 1236
 Db 1021 GGGCAAGGAGCAAGGAGGCTGCTTCCAGGATGGGAGGAGGAGGAGGAGGAGGAGGAG 1080
 QY 1237 GCGAAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1296
 Db 1081 GCGAAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
 QY 1297 AAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1356
 Db 1141 AAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
 QY 1357 CTTGCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1416
 Db 1201 CTTGCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 QY 1417 TCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1476
 Db 1261 TCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 QY 1477 AAGAAGCCCTCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1536

Db 1321 AAGAAGCCCTCAGCCTGGTGTAGTCTCCCTGAGCCTGTGTAAATCTTAAGTCTAAAGAT 1380
 QY 1537 GATGAA 1542
 Db 1381 GATGAA 1386
 RESULT 9
 AX403084
 LOCUS AX403084 1114 bp DNA linear PAT 07-JUN-2002
 DEFINITION Sequence 1 from Patent WO0198355.
 ACCESSION AX403084
 VERSION AX403084.1 GI:21388033
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 Lewis, D., Adams, S.H. and Yu, X.X.
 Cgi-69 compositions and methods of use
 Patent: WO 0198355-A 1 27-DEC-2001;
 Genentech, Inc. (US); Curagen Corporation (US)
 FEATURES
 source
 1..1114
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 194 a 349 c 332 g 239 t
 ORIGIN
 Query Match 68.2%; Score 1054.4; DB 6; Length 1114;
 Best Local Similarity 97.8%; Pred. No. 2.6e-222;
 Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
 QY 105 CTCAAGCTTCAAGATGCTGCTGACGAGCCTCGGGGCAATCGAGCCTCGGGGATCGAGCCTCGGCAATGCT 164
 Db 1 CTCAAGCTTCAAGATGCTGCTGACGAGCCTCGGGGATCGAGCCTCGGGGATCGAGCCTCGGCAATGCT 60
 QY 165 GGCCTCAGGACACCGGGCTGTGGTACCTCTCTTCATGACACCTCGAGCCTCGAGGAGTGA 224
 Db 61 GGCCTCAGGACACCGGGCTGTGGTACCTCTCTTCATGACACCTCGAGCCTCGAGGAGTGA 120
 QY 225 GGTTCGCTGAGTCTCAGCGGCTTCATGCGGCTTCATGCGGAGGCTGATGCTTCCTCCAGCT 284
 Db 121 GGTTCGCTGAGTCTCAGCGGCTTCATGCGGCTTCATGCGGAGGCTGATGCTTCCTCCAGCT 180
 QY 285 GTGGAGCCTCTCTTACCAAT-----GGAAGTGCCTCT 320
 Db 181 GTGGAGCCTCTCTTACCAATTTGCCCTCTCTCTCAATTCACAGGAGTGCCTCT 240
 QY 321 GTATTCATGCTGCTCTGAGGCTCTGTACCTGTGCCAAATGCTGCCGCTGTGCCAC 380
 Db 241 GTATTCATGCTGCTCTGAGGCTCTGTACCTGTGCCAAATGCTGCCGCTGTGCCAC 300
 QY 381 CTGGTTTCAAGACCTTACCGCTTCACTGGCACCATGATGCTTCGTGAAGATCTGAG 440
 Db 301 CTGGTTTCAAGACCTTACCGCTTCACTGGCACCATGATGCTTCGTGAAGATCTGAG 360
 QY 441 GCACGAGGAGCAGGAGCCTCTGGAGGGCTTCGCCACCTCGCTGATGATGATGATGATG 500
 Db 361 GCACGAGGAGCAGGAGCCTCTGGAGGGCTTCGCCACCTCGCTGATGATGATGATGATG 420
 QY 501 AGCTACCGCCATCTACTTCACTGCTTATGACCAATGACGAGGCTTCCTGTGGTGGAGC 560
 Db 421 AGCTACCGCCATCTACTTCACTGCTTATGACCAATGACGAGGCTTCCTGTGGTGGAGC 480
 QY 561 CTTGACTCTGAGCTTACGACCCCATGCTGGTGGCGCTGGCCGCTGGCCGCTGGCCGCTGG 620
 Db 481 CTTGACTCTGAGCTTACGACCCCATGCTGGTGGCGCTGGCCGCTGGCCGCTGGCCGCTGG 540
 QY 621 GACTCTGATGAGGCTTATCGGACCAAGCTCAGGCTCAGCATGTGTCGTA 680

```

Db 541 GACTGTGATCAGCCCCCTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGCTGA 600
QY 681 CCGGAGCTGGGCTGTGTTCGAACCTGACAGTGGCTCAGGCTGGCTGGCTCACTGTG 740
Db 601 CCGGAGCTGGGCTGTGTTCGAACCTGACAGTGGCTCAGGCTGGCTGGCTCACTGTG 660
QY 741 GCTGGGCTGGGCCCCCACTGACCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 800
Db 661 GCTGGGCTGGGCCCCCACTGACCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
QY 801 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGCCGGAAGACACAGACTTCTGTGG 860
Db 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGCCGGAAGACACAGACTTCTGTGG 780
QY 861 CATGAGCTTTGGTGTGTCATCTCAGGACGGTGGCTGCAGTGTGACTCTACCTT 920
Db 781 CATGAGCTTTGGTGTGTCATCTCAGGACGGTGGCTGCAGTGTGACTCTACCTT 840
QY 921 TGACGTGTAAAGACCCCAAGCTGCTTCGGAGAGTGGCTGAGAGTGAAGTGA 980
Db 341 TGACGTGTAAAGACCCCAAGCTGCTTCGGAGAGTGGCTGAGAGTGAAGTGA 900
QY 981 CCCCTGTCATGTGAGCTCCACTGCTGCTGCTGCGGAGATCCGGGCGGAGTGGGAC 1040
Db 901 CCCCTGTCATGTGAGCTCCACTGCTGCTGCGGAGATCCGGGCGGAGTGGGAC 960
QY 1041 CAAGGAGCTTTGAGGCTTCCTTCCTGATCATCAAGCTGCGCCCTCTGTGCCAT 1100
Db 961 CAAGGAGCTTTGAGGCTTCCTTCCTGATCATCAAGCTGCGCCCTCTGTGCCAT 1020
QY 1101 CATGATCAGCACTATGATGCGCAAGAGCTTCTCAGAGGCTGAACAGGAACGGCT 1160
Db 1021 CATGATCAGCACTATGATGCGCAAGAGCTTCTCAGAGGCTGAACAGGAACGGCT 1080
QY 1161 TTGGGCGGCTGAAGGGGCAAGAGGCAAGGAC 1194
Db 1081 TTGGGCGGCTGAAGGGGCAAGAGGCAAGGAC 1114

RESULT 10
AF317711
LOCUS Homo sapiens mitochondrial carrier protein CGI-69 long form mRNA,
DEFINITION complete cds.
ACCESSION AF317711
VERSION AF317711.1 GI:12620399
KEYWORDS "E
SC "E
NISM
REFERENCE 1 (bases 1 to 1114)
AUTHORS Yu,X.X., Lewin,D.A., Zhong,A., Brush,J., Schow,P.W., Sherwood,S.W.,
Pan,G. and Adams,S.H.
TITLE Overexpression of the human 2-oxoglutarate carrier lowers
mitochondrial membrane potential in HEK-293 cells: contrast with
the unique cold-induced mitochondrial carrier CGI-69
JOURNAL Biochem. J. 353 (Pt 2), 369-375 (2001)
MEDLINE 2106063
PUBMED 11139402
REFERENCE 2 (bases 1 to 1114)
AUTHORS Yu,X.X., Lewin,D.A., Zhong,A., Brush,J., Schow,P.W., Sherwood,S.W.,
Pan,G. and Adams,S.H.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2000) Endocrinology, Genentech, Inc., 1 DNA Way,
South San Francisco, CA 94080, USA
FEATURES
source
1. .1114
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue type="liver"
14. .1033
/note="similar to CGI-69 encoded by GenBank Accession
CDS

```

```

Number AF151827"
/codon_start=1
/product="mitochondrial carrier protein CGI-69 long form"
/protein_id="AAG60687.1"
/db_xref="GI:12620400"
/translation="MADQDPAGISPLQOMVAGTGAUVTSLEWTPLDVVKVRLQSORP
SMASLSPSSRLNSLYTKLPSLSASTGKLLYNGVLEPLICPNRGARCAWFDQPT
RFTGMDAFVKIVRHEGRTLMGSLPATVMTVPATAYFYAYDQLKALCGRLTSD
LYAPWALGALRGITVVISPLEMRKTLQAHVSYRELQACVRYTAOQGRSLWLG
WPTALRDVPSALYMYNVELKSWLNGLRPKDQTSVGMFVAGGISGTVAAVLTLPF
DVVKTRQVALGAEVAVRNPLHVDSTWLLLRIRAESGKGLFAGFLRIIKAAPSC
AIMISTVEFGKSFQRLNQRLLGG"
BASE COUNT 194 a 349 c 332 g 239 t
ORIGIN
Query Match 68.2%; Score 1054.4; DB 9; Length 1114;
Best Local Similarity 97.8%; Pred. No. 2.6e-222;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 105 CTGAGCTTCAAGATGGCTGACAGGACCCCTGGGCGCATGCCCTCCAGCAAAATGGT 164
Db 1 CTGAGCTTCAAGATGGCTGACAGGACCCCTGGGCGCATGCCCTCCAGCAAAATGGT 60
QY 165 GGCCTCAGGCACCGGGGCTGTGTTACCTCTCTCTTCATGACACCCCTGGAGCTGGTGA 224
Db 61 GGCCTCAGGCACCGGGGCTGTGTTACCTCTCTCTTCATGACACCCCTGGAGCTGGTGA 120
QY 225 GGTTCGCTGCACTCTCAGCGGCTCCATGGCCAGCGAGTGTGCTTCTCCAGACT 284
Db 121 GGTTCGCTGCACTCTCAGCGGCTCCATGGCCAGCGAGTGTGCTTCTCCAGACT 180
QY 285 GTGAGGCTCTCTATACAAAT-----GGAGTGCCTCT 320
Db 181 GTGAGGCTCTCTATACAAATTTGCCCTCTCTCTCCAAATCCAGGAAAGTGCCTCT 240
QY 321 GTATTGCAATGTGTCTGGAGCTCTGTACCTGTGCCCAATGTGCCCTGTGCCAC 380
Db 241 GTATTGCAATGTGTCTGGAGCTCTGTACCTGTGCCCAATGTGCCCTGTGCCAC 300
QY 381 CTGTTTCAAGACCTACCCGCTCTACTGGCACCATTGGATGCTTCTGTAAGATCTGTAG 440
Db 301 CTGTTTCAAGACCTACCCGCTCTACTGGCACCATTGGATGCTTCTGTAAGATCTGTAG 360
QY 441 GCACGAGGCGACCGAGGACCTCTGGAGCGCTTCCGCCCAATGTGCCCTGTGCCAC 500
Db 361 GCACGAGGCGACCGAGGACCTCTGGAGCGCTTCCGCCCAATGTGCCCTGTGCCAC 420
QY 501 AGCTACCGCCATCTACTCTACTGCTATGACCACTGAAGGCTTCTCTGTGTGGTGGAC 560
Db 421 AGCTACCGCCATCTACTCTACTGCTATGACCACTGAAGGCTTCTCTGTGTGGTGGAC 480
QY 561 CTGACCTCTGACCTCTACGCCACCCATGGTGGCTGGCGCTGGCGCTGGGACCGT 620
Db 481 CCGTACCTCTGACCTCTACGCCACCCATGGTGGCTGGCGCTGGCGCTGGGACCGT 540
QY 621 GACTGTGATCAGCCCTCTGAGCTTATGGGACAAAGCTGAGGCTCAGAGTGTCTCTA 680
Db 541 GACTGTGATCAGCCCTCTGAGCTTATGGGACAAAGCTGAGGCTCAGAGTGTCTCTA 600
QY 681 CCGGAGCTGGGCTGTGTTCTGAACTGCACTGGCTCAGGCTGGCTGGCTCACTGTG 740
Db 601 CCGGAGCTGGGCTGTGTTCTGAACTGCACTGGCTCAGGCTGGCTGGCTCACTGTG 660
QY 741 GCTGGCTGGGCGCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 800
Db 661 GCTGGCTGGGCGCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
QY 801 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGCCGGAAGACACAGACTTCTGTGG 860
Db 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGCCGGAAGACACAGACTTCTGTGG 780
QY 861 CATGAGCTTTGGTGTGTCATCTCAGGACGGTGGCTGCAGTGTGACTCTACCTT 920

```


RESULT 12
AC124203
LOCUS

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYVK

1	1426:	contig	of 1426 bp in length
1427	1526:	gap	of unknown length
1527	2912:	contig	of 1386 bp in length
2913	3012:	gap	of unknown length
3013	4151:	contig	of 1139 bp in length
4152	4251:	gap	of unknown length
4252	5629:	contig	of 1378 bp in length
5630	5723:	gap	of unknown length
5730	6926:	contig	of 1197 bp in length
6927	7026:	gap	of unknown length
7027	8163:	contig	of 1137 bp in length
8164	8263:	gap	of unknown length
8264	9582:	contig	of 1319 bp in length
9583	9682:	gap	of unknown length
9683	11389:	contig	of 1707 bp in length
11390	11489:	gap	of unknown length
11490	13054:	contig	of 1565 bp in length
13055	13154:	gap	of unknown length
13155	14410:	contig	of 1256 bp in length
14411	14510:	gap	of unknown length
14511	15590:	contig	of 1080 bp in length
15591	15690:	gap	of unknown length
15691	16982:	contig	of 1292 bp in length
16983	17082:	gap	of unknown length
17083	18917:	contig	of 1835 bp in length
18918	19017:	gap	of unknown length
19018	20730:	contig	of 1713 bp in length
20731	20830:	gap	of unknown length
20831	22139:	contig	of 1309 bp in length
22140	22239:	gap	of unknown length
22240	23606:	contig	of 1367 bp in length
23607	23706:	gap	of unknown length
23707	25243:	contig	of 1537 bp in length
25244	25343:	gap	of unknown length
25344	26593:	contig	of 1250 bp in length
26594	26693:	gap	of unknown length
26694	28414:	contig	of 1721 bp in length
28415	28514:	gap	of unknown length
28515	31435:	contig	of 2921 bp in length
31436	31535:	gap	of unknown length
31536	33531:	contig	of 1996 bp in length
33532	33631:	gap	of unknown length
33632	36326:	contig	of 2695 bp in length
36327	36426:	gap	of unknown length
36427	40073:	contig	of 3647 bp in length
40074	40173:	gap	of unknown length
40174	42131:	contig	of 1958 bp in length
42132	42231:	gap	of unknown length
42232	44453:	contig	of 2222 bp in length
44454	44553:	gap	of unknown length
44554	46183:	contig	of 1630 bp in length
46184	46283:	gap	of unknown length
46284	48384:	contig	of 2101 bp in length
48385	48484:	gap	of unknown length

	*	48485	52244:	contig of 3760 bp in length
	*	52245	52344:	gap of unknown length
	*	52345	54687:	contig of 2343 bp in length
	*	54688	54787:	gap of unknown length
	*	54788	57211:	contig of 2424 bp in length
	*	57211	57311:	gap of unknown length
	*	57312	59196:	contig of 1887 bp in length
	*	59199	59298:	gap of unknown length
	*	59299	61525:	contig of 2227 bp in length
	*	61526	61625:	gap of unknown length
	*	61626	64137:	contig of 2512 bp in length
	*	64138	64237:	gap of unknown length
	*	64238	66994:	contig of 2757 bp in length
	*	66995	67094:	gap of unknown length
	*	67095	69482:	contig of 2388 bp in length
	*	69483	69592:	gap of unknown length
	*	69583	72067:	contig of 2485 bp in length
	*	72068	72167:	gap of unknown length
	*	72168	74228:	contig of 2061 bp in length
	*	74229	74328:	gap of unknown length
	*	74329	78279:	contig of 3951 bp in length
	*	78280	78379:	gap of unknown length
	*	78380	81948:	contig of 3569 bp in length
	*	81949	82048:	gap of unknown length
	*	82049	85828:	contig of 3780 bp in length
	*	85829	85928:	gap of unknown length
	*	85929	90452:	contig of 4524 bp in length
	*	90453	90552:	gap of unknown length
	*	90553	93607:	contig of 3055 bp in length
	*	93608	93707:	gap of unknown length
	*	93708	96753:	contig of 3046 bp in length
	*	96754	96853:	gap of unknown length
	*	96854	99186:	contig of 2333 bp in length
	*	99187	99286:	gap of unknown length
	*	99287	103012:	contig of 3726 bp in length
	*	103013	103112:	gap of unknown length
	*	103113	107318:	contig of 4206 bp in length
	*	107319	107418:	gap of unknown length
	*	107419	112328:	contig of 4910 bp in length
	*	112329	112428:	gap of unknown length
	*	112429	117734:	contig of 5306 bp in length
	*	117735	117834:	gap of unknown length
	*	117835	121612:	contig of 3778 bp in length
	*	121613	121712:	gap of unknown length
	*	121713	125712:	contig of 4000 bp in length
	*	125713	125812:	gap of unknown length
	*	125813	130929:	contig of 5117 bp in length
	*	130930	131029:	gap of unknown length
	*	131030	133655:	contig of 4626 bp in length
	*	133656	135755:	gap of unknown length
	*	135756	140212:	contig of 4457 bp in length
	*	140213	140312:	gap of unknown length
	*	140313	143560:	contig of 3248 bp in length
Query Match				
Best Local Similarity 35.7%; Score 551.2; DB 2; Length 195701;				
Matches 663; Conservative 83.3%; Pred. No. 2e-11l;				
Mismatches 0; Mismatches 128; Indels 5; Gaps 3;				
QY	382	TGGTTTCAGACCTTACC	GCGTTCACTGGGACCATGTGATGCCTTGGAAGATCGTGAGG	441
Db	72766	TGGTTTTCAGGATCCTACAC	GCGTTCACTGGCACCTTGGATGCCCTCGTGAAGATTGTGAGG	72825
QY	442	CACGAGGGACACGAGGAC	CCCTCTGGAGGGCGCTCCCGGACCCCTGGTGACTGTGCA	501
Db	72826	CAGAGGGACACAGAGAC	CTGTGAGGGCGCTCCGAGCCACCCCTGGTGACTGTGCA	72885
QY	502	GCTACGCCATCTACTT	CACCTGATGACCAACTGAAGGCCCTCTGTGTGTGCGAGCC	561
Db	72886	GCCACTGCCATCTACTT	CACCTGATGATGACCAACTGAGGCCCTCTGTGTGTGCGAGCC	72945
QY	562	CTGACCTCTGACCTCT	ACGACCATGTGGTGGCGCGCTGGCGCCCTCGGSCACCGTG	621
Db	72946	TTGACCTCTGACCTCT	ATGSCAAATGTGGTGGTGGCGCTCGCCCAATGSGCACTGTG	73005

REFERENCE
AUTHORS

1. (Bases 1 to 156175)

Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Ari,J.R., Ayelle,M., Banks,T., Barbara,J., Benton,J., Bimago,J., Bimago,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,K., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinth,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Galis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

```

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GMCE
Center clone name: CH230-161G12
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 99973 bases at least Q40
Consensus quality: 106615 bases at least Q30
Consensus quality: 111100 bases at least Q20

```

	1025: contig	of 1025	bp in length
1026	gap	of unknown	length
1028	contig	of 1093	bp in length
2219	gap	of unknown	length
3219	contig	of 1136	bp in length
3455	gap	of unknown	length
3555	contig	of 1136	bp in length
4691	gap	of unknown	length
4791	contig	of 1165	bp in length

```

* 5956 6055: gap of unknown length
* 6056 7439: contig of 1384 bp in length
* 7440 7539: gap of unknown length
* 7540 8707: contig of 1168 bp in length
* 8708 8807: gap of unknown length
* 8808 10033: contig of 1226 bp in length
* 10034 10133: gap of unknown length
* 10134 11778: contig of 1645 bp in length
* 11779 11878: gap of unknown length
* 11879 13633: contig of 1755 bp in length
* 13634 13733: gap of unknown length
* 13734 14947: contig of 1214 bp in length
* 14948 15047: gap of unknown length
* 15048 16569: contig of 1522 bp in length
* 16570 16669: gap of unknown length
* 16670 17886: contig of 1117 bp in length
* 17887 17886: gap of unknown length
* 17888 19738: contig of 1852 bp in length
* 19739 19838: gap of unknown length
* 19839 21532: contig of 1694 bp in length
* 21533 21632: gap of unknown length
* 21633 22715: contig of 1083 bp in length
* 22716 22815: gap of unknown length
* 22816 24052: contig of 1237 bp in length
* 24053 24152: gap of unknown length
* 24153 25250: contig of 1098 bp in length
* 25251 25350: gap of unknown length
* 25351 26944: contig of 1594 bp in length
* 26945 27044: gap of unknown length
* 27045 28474: contig of 1430 bp in length
* 28475 28574: gap of unknown length
* 28575 29925: contig of 1351 bp in length
* 29926 30025: gap of unknown length
* 30026 31402: contig of 1377 bp in length
* 31403 31502: gap of unknown length
* 31503 33645: contig of 2143 bp in length
* 33646 33745: gap of unknown length
* 33746 35290: contig of 1545 bp in length
* 35291 35390: gap of unknown length
* 35391 36772: contig of 1282 bp in length
* 36773 38565: gap of unknown length
* 38566 40077: contig of 1793 bp in length
* 40078 40177: gap of unknown length
* 40178 41278: contig of 1101 bp in length
* 41279 41378: gap of unknown length
* 41379 42846: contig of 1468 bp in length
* 42847 42946: gap of unknown length
* 42947 44779: contig of 1833 bp in length
* 44780 44879: gap of unknown length
* 44880 47030: contig of 2151 bp in length
* 47031 48446: contig of 1716 bp in length
* 48447 48946: gap of unknown length
* 48947 50071: contig of 1125 bp in length
* 50072 50171: gap of unknown length
* 50172 51617: contig of 1446 bp in length
* 51618 51717: gap of unknown length
* 51718 53821: contig of 2104 bp in length
* 53822 53921: gap of unknown length
* 53922 55611: contig of 1690 bp in length
* 55612 55711: gap of unknown length
* 55712 57525: contig of 1814 bp in length
* 57526 57625: gap of unknown length
* 57626 59106: contig of 1481 bp in length
* 59107 59206: gap of unknown length
* 59207 61142: contig of 1936 bp in length
* 61143 61242: gap of unknown length
* 61243 64511: contig of 3169 bp in length
* 64512 64512: gap of unknown length
* 64512 66886: contig of 2175 bp in length
* 66887 66786: gap of unknown length

```

```

* 66787 68966: contig of 2180 bp in length
* 68967 69066: gap of unknown length
* 69067 72130: contig of 3064 bp in length
* 72131 72230: gap of unknown length
* 72231 73778: contig of 1548 bp in length
* 73779 73878: gap of unknown length
* 73879 76273: contig of 2395 bp in length
* 76274 76373: gap of unknown length
* 76374 78932: contig of 2559 bp in length
* 78933 79032: gap of unknown length
* 79033 81984: contig of 2951 bp in length
* 81984 82083: gap of unknown length
* 82084 84716: contig of 2633 bp in length
* 84717 84816: gap of unknown length
* 84817 87801: contig of 2984 bp in length
* 87801 87900: gap of unknown length
* 87901 92353: contig of 4453 bp in length
* 92354 92453: gap of unknown length
* 92454 96499: contig of 4046 bp in length
* 96500 96599: gap of unknown length
* 96600 98503: contig of 2904 bp in length
* 98504 99603: gap of unknown length
* 99604 103250: contig of 3647 bp in length
* 103251 103350: gap of unknown length
* 103351 106405: contig of 3055 bp in length

Query Match      34.7%; Score 537.2; DB 2; Length 156175;
Best Local Similarity 77.6%; Pred. No. 2.5e-108;
Matches 709; Conservative 0; Mismatches 183; Indels 22; Gaps 4;

QY 263 AGCTGATGCTTCTCCAGAGCTGGAGCTCTCTCTATACCAATGAAGTGCCTCTCT 322
Db 77061 AGATTCTGAGTCTCTCTACACCTATACACCTCTGCTCTCGGGAAGGCTCTCAT 77120

QY 323 ATTGCATGCTGCTGGAGCTCTGTACTGTGCCAAATGTTGCCGCTGTCACCT 382
Db 77121 ACTGCTATGCTGCTGGAGCTCTGGAGCTGTGCCAAATGATACCTGCTGCGCACCT 77180

QY 383 GGTTCACAGACCTACCGCTTCTACCTGGCACCATGATGCTTCTGTAAGATGTAAGC 442
Db 77181 GGTTCAGATCTTACAGGTTTCTGGCACCCTTGGATGCTTGTGAAGATGTAAGC 77240

QY 443 AGAGGGCACCAGACCTCTGGAGGGCTCTCCCGCCACCTGCTGGTGTGATGTCGACG 502
Db 77241 -CGAGGGCACCAGACCTCTGGAGGGCTCTCCCGCCACCTGCTGGTGTGATGTCGACG 77299

QY 503 CTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGTTGTGTCGAGGCC 562
Db 77300 CTACCGCATCTATTTCACTGCTATGACCACTGAAGGCTTCTGTTGTGTCGAGGCC 77355

QY 563 TGACCTCTGACCTCTACGACCCCTGCTGGTGGGGCTGCGCCGCTGCGCCTGCGACCTGA 622
Db 77356 TGACCTCTGACCTCTACGACCCCTGCTGGTGGGGCTGCGCCGCTGCGCCTGCGACCTGA 77415

QY 623 CTGTGATCAGCCCTCTGAGCTTATGCGGACAAAGCTGACAGCTGTCGCTACC 682
Db 77416 CAGTGTGAGCCCTCTGAGCTTATGCGGACAAAGCTGACAGCTGTCGCTACC 77475

QY 683 GGGAGCTGGGCTGCTGTTTCAACTGCTGAGTGGCTCAGGCTGCTGCTGCTGCTGCTG 742
Db 77476 GCAAACTGGCTGCTGTTTCAACTGCTGAGTGGCTCAGGCTGCTGCTGCTGCTGCTG 77535

QY 743 TGGGCTGGGGCTGCTGCTGCTTCTGAGATGTCCTTCTCAGCCTGCTGCTGCTCACT 802
Db 77536 TGGGCTGGGGCTTCTGAGTCTTCTGAGATGTCCTTCTGAGCCTGCTGCTGCTCACT 77595

QY 803 ATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTCAAGCCGAGGACCACTTCTGTCGCA 862
Db 77596 ACAAGCTGGTGAAGGCTGAGTGAATGGGCTGAGGACCACTTCTGTCGCA 77655

QY 863 TGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922
Db 77656 TCAGTTTGTGCTTAATGGCTTCTGAGGATGGGGCTGCTGCTGCTGCTGCTGCTGCTG 77715

```

QY	923	ACGTGTAAGAACCACCGCAGGTGCTCTGGAGGAGTGGAGGCTGTGAGAGTGAACC	982
Db	77716	ATGTGATGAAGACACAGGACAATGTCACTGTGAGCGGTGGAGGCTATGAGA	77768
QY	983	CCCTGCATGTGGAATCCCACTGGCTGTGCTGCGAGGATCCGGGCGGAGTCGGGCACCA	1042
Db	77769	-----GTTGCGCTCCACCTGGCTCTGCTTCGGAGAAATCCAGGCTGAAATCGGCACCA	77820
QY	1043	AGGGACTCTTTGCGAGGCTTCCTTCTCCGAGATCAAGGCTGCCCTCTCTGTGCCATCA	1102
Db	77821	GGGGAATC--TGCAGGTTTCTCCCGAGATCATCAAGGCTGCACCTCTGTGCCATCA	77878
QY	1103	TGATCAGCACCTATGAGTTTCGGCAAAAGCTTCTTCCAGAGGCTGAACACGAGCAGCGGCTTC	1162
Db	77879	GCAGCAGCACCTATGAGTTTGGCAAAAGCTTCTTCCAGAGGCTGAACCAAAAGCAGCTTC	77938
QY	1163	TGGGCGGCTGAAAG	1176
Db	7939	TGGGCGCACTGAAAG	77952
RESULT	14		
LOCUS	AC123139/c		
DEFINITION	Rattus norvegicus clone CH230-11K3, *** SEQUENCING IN PROGRESS ***;	128802 bp	DNA linear HTG 13-JUL-2002
	29 unordered pieces.		
ACCESSION	AC123139		
VERSION	AC123139.2		
KEYWORDS	HTG; HTGS PHASE1.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		

Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 128802)
2 (bases 1 to 128802)
Worley,K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 128802)
Worley,K.C.
Direct Submission
Submitted (13-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 3, 2002 this sequence version replaced gi:21239891.
----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GMFG
Center Clone name: CH230-11K3
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 108533 bases at least Q40
Consensus quality: 111720 bases at least Q30
Consensus quality: 113548 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length.
 (see http://www.hgsc.bcm.tmc.edu/docs/GenBank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 29 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1	1434:	contig of 1434 bp in length
1435	1534:	gap of unknown length
1535	3218:	contig of 1684 bp in length
3219	3318:	gap of unknown length
3319	4531:	contig of 1213 bp in length
4532	4631:	gap of unknown length
4632	5802:	contig of 1171 bp in length
5803	5902:	gap of unknown length
5903	7060:	contig of 1158 bp in length
7061	7160:	gap of unknown length
7161	8514:	contig of 1354 bp in length
8515	8614:	gap of unknown length
8615	9672:	contig of 1058 bp in length
9673	9772:	gap of unknown length
9773	12180:	contig of 2408 bp in length
12181	12280:	gap of unknown length
12281	13463:	contig of 1183 bp in length
13464	13563:	gap of unknown length
13564	16106:	contig of 2543 bp in length
16107	16206:	gap of unknown length
16207	18984:	contig of 2678 bp in length
18985	18984:	gap of unknown length
18985	20532:	contig of 1548 bp in length
20533	20632:	gap of unknown length
20633	23210:	contig of 2578 bp in length
23211	23310:	gap of unknown length
23311	25515:	contig of 2205 bp in length
25516	25615:	gap of unknown length
25616	27411:	contig of 1796 bp in length

Query Match	34.3%	Score 530	DB 2	Length 128802
Best Local Similarity	75.9%	Prod. No. le-106		
Matches 733	Conservative	0	Mismatches 200	Indels 33
Gaps	5			
QY 50	TGGGCGCGCTGGCCCTGGCGCGCGGC	CAGGACCA	CAGCCTAGAGCCAGGACTGA	109
DB 78125				
QY 110	GCTTCAAGATGGCTGACACGACCTTGGGGCATCAGCCCTCCAGCAAAATGTTGGCTT	AACGACCGGCTCAGAGCCAGGATTGA	78066	
DB 78065				
QY 170	CAGGACCGGGGCTTGGTTACTCTCTTCATGACACCCCTGGACCTGGTGAAGGTTT		229	
DB 78005				
QY 230	GCCTCAGCTTCA-----GGGCGCTCCATGGCCAGGAGCTGATCCCTTCCTCCAGACT		284	
DB 77945				
QY 285	GTGGAGCTCTTCTTATACAAAT-----GGAAGTGGCTCT		320	
DB 77885				
QY 321	GTATTGCATGGTGTCTCGAGCTCTGTACTGTGCCAAATGGTCCCGCTGTGCCAC		380	
DB 77825				
QY 381	CTGGTTTCAAGACCTTACCCTTCTCACTGGCACCATTGATGCCTTCGTGAGATCGTAG		440	
DB 77765				
QY 441	GCACGAGGACACGAGCCCTCTGGAGCGGCTCCCGGCCACCTGGTGTGATGACTGTGCC		500	
DB 77705				
QY 501	AGCTACCGCCATCTACTTCACTGCTATGACAACTGAAGGCTTCTCTGTGTGGTAGC		560	
DB 77646				

```

Db 121 TGAACACAGACCGGCTTCTGGGGCTGAAGGGGCAAGGAGCAAGCCCGTCTCTC 180
QY 1205 CACCGGATGGGAGAGGGCAGGAGAGACCCAGCCAGTGCTTTTCTCAGCACTGAGG 1264
Db 181 CCACGGATGGGAGAGGGCAGGAGAGACCCAGCCAGTGCTTTTCTCAGCACTGAGG 240
QY 1265 GAGGGGGCTTCTTTCCCTTCCCTCCCGGCGACAAGCTCCAGGGCAGGGCTGTCCCTCTGG 1324
Db 241 GAGGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGGCAGGGCTGTCCCTCTGG 300
QY 1325 GCGGCCAGCACTTCCCTCAGACACAACCTTCTCTGCTGCTCCAGTCGTGGGATCATCA 1384
Db 301 GCGGCCAGCACTTCCCTCAGACACAACCTTCTCTGCTGCTCCAGTCGTGGGATCATCA 360
QY 1385 CTTACCCACCCCGGCAAGTTCAAGACCAAAATCTCCAGCTGCCCCCTTCTGTTCCTGT 1444
Db 361 CTTACCCACCCCGGCAAGTTCAAGACCAAAATCTCCAGCTGCCCCCTTCTGTTCCTGT 420
QY 445 GTTGTCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 1504
Db 421 GTTGTCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 480
QY 1505 CTGACCCCTTCTTAATTCCTTAAGTCTTAAGATGATGAATTC 1546
Db 481 CTGACCCCTTCTTAATTCCTTAAGTCTTAAGATGATGAATTC 522

```

Search completed: January 22, 2003, 23:26:32
 Job time : 4855.54 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compu

Copyright (c) 1993 - 2003 CompuGen Ltd.
GenCore version 3.1.3

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:31:07 ; Search time 354.534 Seconds
(without alignments)
9820.189 Million cell updates/sec

Title: US-09-888-358-2

Perfect score: 1546

Sequence: 1 ggctaggctgcgctgcgagcg.....gtctaaagatgatgaacttc 1546

Scoring table: IDENTITY NUC

```

SCROLLING LABEL: IDENTIFY_NOC
Gapop 10.0 , Gapext 1.0

```

Searched: 2185239 seqs. 1125999159 residues

Tc: number of hits satisfying chosen parameters: 4370478

Minimum DB seq Length: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 200000000
```

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Paraphase : M Consona 101003. *

```

Database :
1: /SID32/cgcdata/genseq/genseq-emb1/NA1980.DAT.*
2: /SID32/cgcdata/genseq/genseq-emb1/NA1981.DAT.*
3: /SID32/cgcdata/genseq/genseq-emb1/NA1982.DAT.*
4: /SID32/cgcdata/genseq/genseq-emb1/NA1983.DAT.*
5: /SID32/cgcdata/genseq/genseq-emb1/NA1984.DAT.*
6: /SID32/cgcdata/genseq/genseq-emb1/NA1985.DAT.*
7: /SID32/cgcdata/genseq/genseq-emb1/NA1986.DAT.*
8: /SID32/cgcdata/genseq/genseq-emb1/NA1987.DAT.*
9: /SID32/cgcdata/genseq/genseq-emb1/NA1988.DAT.*
10: /SID32/cgcdata/genseq/genseq-emb1/NA1989.DAT.*
11: /SID32/cgcdata/genseq/genseq-emb1/NA1990.DAT.*
12: /SID32/cgcdata/genseq/genseq-emb1/NA1991.DAT.*
13: /SID32/cgcdata/genseq/genseq-emb1/NA1992.DAT.*
14: /SID32/cgcdata/genseq/genseq-emb1/NA1993.DAT.*
15: /SID32/cgcdata/genseq/genseq-emb1/NA1994.DAT.*
16: /SID32/cgcdata/genseq/genseq-emb1/NA1995.DAT.*
17: /SID32/cgcdata/genseq/genseq-emb1/NA1996.DAT.*
18: /SID32/cgcdata/genseq/genseq-emb1/NA1997.DAT.*
19: /SID32/cgcdata/genseq/genseq-emb1/NA1998.DAT.*
20: /SID32/cgcdata/genseq/genseq-emb1/NA1999.DAT.*
21: /SID32/cgcdata/genseq/genseq-emb1/NA2000.DAT.*
22: /SID32/cgcdata/genseq/genseq-emb1/NA2001A.DAT.*
23: /SID32/cgcdata/genseq/genseq-emb1/NA2001B.DAT.*
24: /SID32/cgcdata/genseq/genseq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match %	Length	DB	ID	Description
1	1546	100.0	1546	24	AA171020	Human mitochondria
2	1500.4	97.1	1560	21	AAAC90452	Human uncoupling p
3	1421.4	91.9	1662	21	AAAG76992	Human ORFX ORF2347
4	1054.4	68.2	1114	24	AA171019	Human mitochondria
5	522	33.8	537	21	AAAG06806	Human immunogenic
6	522	33.8	537	22	AAAG38151	Human prostate cdn
7	522	33.8	537	22	AAAG93722	Human prostate-spe
8	522	33.8	537	22	AAAH85036	Human prostate-spe
9	522	33.8	537	22	AAAG02787	Prostate tumour an

10	522	33.8	537	24	ABL95181
11	517.2	33.5	2170	24	ABK347435
12	470.4	30.4	704	24	ABK199647
C 13	445.4	28.8	452	24	ABN946435
C 14	437.4	28.3	788	24	ABL01419
C 15	242.2	15.7	6712	21	PAC766691
C 16	218.4	14.1	240	24	ABN146481
C 17	185.8	12.0	453	24	ABN481508
18	172	11.1	275	22	ABL005555
19	172	11.1	275	23	ABK72077
C 20	172	11.1	275	24	ABK91669
C 21	170.2	11.0	512	24	ABK62422
C 22	166	10.7	174	16	ABL26018
C 23	151.4	9.8	1258	24	ABL89865
C 24	142.6	9.2	397	21	AAH30252
C 25	128.8	8.3	1449	23	ABL04893
C 26	128.8	8.3	3928	23	ABL04892
C 27	122	7.9	422	22	ABK53810
C 28	111.8	7.2	145	23	ABL24583
C 29	79.6	5.1	3645	23	ABL24582
C 30	77.2	5.1	1529	23	ABL24581
C 31	77.4	5.0	2516	24	ASL90270
C 32	74.8	4.8	452	22	ABA59262
C 33	74.8	4.8	452	22	ABK07474
C 34	74.8	4.8	432	22	AKK33259
C 35	74.8	4.8	432	22	ABL33053
C 36	74.8	4.8	432	24	ABSO8091
C 37	74.8	4.8	496	22	ABA59322
C 38	74.8	4.8	496	22	ABK07540
C 39	74.8	4.8	496	22	AAK33336
C 40	74.8	4.8	496	22	ABL39123
C 41	74.8	4.8	486	24	ABSO8171
C 42	63.8	4.1	367	24	ABL01420
C 43	61	3.9	997	21	AAZ61628
C 44	61	3.9	997	22	AAC99561
C 45	61	3.9	997	24	ABL34713

ALIGNMENTS

RESULT 1

RESUL1	
AAI71020	
ID	AAI71020 standard; cDNA; 1546 BP.
XX	
AC	AAI71020;
XX	
DT	18-MAR-2002 (first entry)
XX	
DE	Human mitochondrial carrier protein CGI-69 cDNA.
XX	
XX	CGI-69; mitochondrial carrier protein; human; metabolic disease; obesity; cachexia; tumour; cancer; infection; immunomodulator; antitumour; virucide; antibacterial; anorectic; antidiabetic; brown adipose tissue; diagnosis; gene therapy; ss.
XX	

Key	Location/Qualifiers
-----	---------------------

Key	Location/Quantity
CDS	118..1173

```

CDS
148,::11/3
/*taq= a

```

XXIX

WO200198355-A2.

XX
NY 0400198333-AZ.

27-DEC-2001.

27-DEC-2001.

22-JUN-2001: 200

XX
FF
22-JUN-2001; 2000

22-JUN-2000: 200

22-JUN-2000; 2000XX

XX
PA
(GETH) GENENTEC

PA (GETH) GENENTEC
PA (CIRA-) CUPAGEN

(CURA-) CURAGEN

Q

Lewin D, Adams SH, Yu XX;

WPI: 2002-114569/15.

CGI-69 polypeptides and polynucleotides useful for treating metabolic disorders including cachexia, obesity, diabetes and cancers -
Disclosure; Page 11-12; 107pp; English.

Disclosure; Page 11-12; 107pp; English.

The present sequence is that of cDNA encoding human CGI-69 (see UAM0507R). Various CGI-69 clones were isolated from human liver by PCR amplification and cloning. Some diverged from the previously known CGI-69 sequence in that they encoded an 8-amino acid insert preceded by a W44L change, and were termed the 'long version' isoform or CGI-69L (see AAI71019). The invention relates to a novel characterization of CGI-69 as a mitochondrial carrier protein, the discovery of CGI-69L, and the discovery of the alteration of the mitochondrial membrane potential through overexpression of carboxyl-FLAG-tagged CGI-69. CGI-69 (including splice variant) nucleic acids and proteins are useful for diagnosing and treating metabolic diseases in humans, including obesity, cachexia and diabetes. Evidence for an important function for CGI-69 in modifying mitochondrial membrane potential in brown adipose tissue (BAT) is presented. The mouse orthologue of CGI-69 is up-regulated in cold-treated BAT. CGI-69 may be involved in cellular thermogenic uncoupling and, therefore, may be used to diagnose and treat specific perturbations in metabolic pathways. Altering the expression of CGI-69 through gene therapy provides a means of treating metabolic diseases, such as obesity or cachexia, or of increasing or decreasing body weight. Decreasing the activity of CGI-69 can be used to treat cachexia, tumours, cancers, viral infections and bacterial infections, while increasing its activity can be used to treat obesity, tumours, cancers, viral infections and bacterial infections (all claimed). Methods are also claimed for determining whether a compound up-regulates or down-regulates expression of a CGI-69 gene, and of screening for a mutation in the CGI-69 gene.

Sequence 1546 BP; 272 A; 489 C; 459 G; 326 T; 0 other;

Query Match	100.0%; Score 1546; DB 24; Length 1546;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1546; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1	GCCTAGGTGGCGTCCGAGCGCGCGAGGACGAGCGTAAATGGCGCGCGCT 50
DB 1	GGCTAGGTGGCGTCCGAGCGCGCGAGGACGAGCGTAAATGGCGCGCGCT 50
QY 61	GGCCCTGGGCGCGCGCGCGAGCAGCAGCTTAGAGCTTCAAGATG 120
DB 61	GGCCCTGGGCGCGCGCGCGAGCAGCAGCTTAGAGCTTCAAGATG 120
QY 121	GCTGACACAGGACCTCGGGGCACTAGCCCCCTCCAGCAATGCTGGCCTCAGGCAACGGG 180
DB 121	GCTGACACAGGACCTCGGGGCACTAGCCCCCTCCAGCAATGCTGGCCTCAGGCAACGGG 180
QY 181	GCTGTGTTTACTCTCTCTTCAAGCAGCCCCCTGAGCGTGGTGAAGGTTGCGCTGCACTCT 240
DB 181	GCTGTGTTTACTCTCTCTTCAAGCAGCCCCCTGAGCGTGGTGAAGGTTGCGCTGCACTCT 240
QY 241	CAGCGGCCCTCCATGSCCAGCGAGCTGATGCTTCCTCCAGACTGTGAGAGCTCTCCTAT 300
DB 241	CAGCGGCCCTCCATGSCCAGCGAGCTGATGCTTCCTCCAGACTGTGAGAGCTCTCCTAT 300
QY 301	ACCAAATGGAAGTGCCTCCTGTATTGCAATGGTGTCTTGAGSCCTCTCTACCTGTGCCCA 360
DB 301	ACCAAATGGAAGTGCCTCCTGTATTGCAATGGTGTCTTGAGSCCTCTCTACCTGTGCCCA 360
QY 361	AATGTGCCCCGTGTGCCACCTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATTGAT 420
DB 361	AATGTGCCCCGTGTGCCACCTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATTGAT 420
QY 421	GCCTTCTGTGAAGATCTGTGAGGCACGAGGGCACAGGACCTCTCTGGAGCGGCTCCCGGCC 480

Db 1501 CTCCCTGACCTTGTTAATTCCTTAAGCTAAAGATGATGAACCTTC 1546

RESULT 2

AAC90452

ID	AAC90452 standard; cDNA; 1560 BP.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

AAC90452;

12-MAR-2001 (first entry)

Human uncoupling protein cDNA #1.

Human; uncoupling protein; immunosuppressive; antiarthritis;
antirheumatic; antiproliferative; cardiac; vasotropic;
cerebroprotective; neuroprotective, antibacterial; ophthalmological;
gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;
gene therapy; cancer; wound; infectious disease, thrombosis; arthritis;
infertility; ss.

omo sapiens.

WO2000061614-A2

19-OCT-2000

06-APR-2000:

09-APR-1999 0918-0138701

08-JUL-1999; 99US-0142821.
18-AUG-1999; 99US-0140418

12-NOV-1999; 99US-0164751.

XX (HUMA-) HUMAN GENOME SCI INC.

XX
I
RUBELL SM, NI J, KOMATSOUKIS

REF: 2000-656322/63.
P-PSDB: AAB50378

Uncoupling proteins and nucleic acid metabolism

IT for detecting, preventing, and curing the immune system's cardiac

Claim 1; Page 303

The present sequence is one of eighteen isolated nucleotide sequences encoding uncoupling proteins. The nucleotide sequences may be used for the detection of various disorders such as cancer, for chromosome identification, as chromosome markers and for numerous other diagnostic or research purposes. The uncoupling protein encoded by the nucleotide sequence may be used to treat disorders such as neural, immune, muscular, reproductive, and endocrine disorders.

renal and proliferative disorders, wounds, infectious diseases, thrombosis, arthritis, and infertility.

Sequence 1560 BP; 286 A; 492 C; 459 G; 323 T; 0 other:

Query Match 97.1%; Score 1500.4; DB 21; Length 1560;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1510; Conservative 0; Mismatches 19; Indels 0

14 GCGAGCGCGGGAGGCCGACGCCTTCC

[illegible]

74 GCGCCGACGAGCACTAGCCTTAGGCCAATGGTTCCTTT

76 GCGCCGACGAGCAGCAGCCTAGCCACCCCTCAATCCCTT

134 CTGCGGGCATCAGCCCCCTCCAGCAATACTCTCCCTCCCT

136 CTGCGGGCATCAGCCCCCTCCAGCAATCTCCCTCCCCTCCAGTCCGC

QY	1274	TGTTTCCCTTCCTCCCGCGCAGAAAGTCTCAGGGCAGGGCTGTCCCTCTGGGCGGCCAG	1339
Db	1276	TGTTTCCCTTCCTCCCGCGCAGAAAGTCTCAGGGCAGGGCTGTCCCTCTGGGCGGCCAG	1339
QY	1334	CACCTTCCTCAGACACAACCTTCCTCCTGCTCCTCCAGTCGTGGGATCATCACTTACCCAC	1393
Db	1336	CACCTTCCTCAGACACAACCTTCCTCCTGCTCCTCCAGTCGTGGGATCATCACTTACCCAC	1395
QY	1394	CCCCCAAGTTCAGACCAAAATCTTCCAGCTGCCCTTCGTTGTTCCCTGTGTTGCTGT	1453
Db	1396	CCCCCAAGTTCAGACCAAAATCTTCCAGCTGCCCTTCGTTGTTCCCTGTGTTGCTGT	1455
QY	1454	AGCTGGGCATGTCTCCAGAACCAAGAACGCCCTCAGCCCTGCTGTAGTCTCCCTGACCCCT	1513
Db	1456	AGCTGGGCATGTCTCCAGAACCAAGAACGCCCTCAGCCCTGCTGTAGTCTCCCTGACCCCT	1515
QY	1514	GTTAAATTCCTTAAGTCTTAAAGATGAA	1542
Db	1516	GTTAAATTCCTTAAGTCTTAAAGATGAA	1544
RESULT	3		
AACT76992			
ID	AACT76992	standard; cDNA; 1662 bp.	
XX			
AC	AACT76992:		
XX			
DT	08-FEB-2001	(first entry)	
XX			
DE	Human ORFX ORF2547	polynucleotide sequence SEQ ID NO:5093.	
XX			
KW	Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;		
KW	vulnary; antipsoriatic; antiparkinsonian; neutropic; neuroprotective;		
KW	anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiac		
KW	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;		
KW	hypotensive; dermatological; immunosuppressive; antiinflammatory;		
KW	antiviral; antibacterial; antifungal; antineumatic; antithyroid;		
KW	antineuemic; gene therapy; cancer; proliferative disorder; hypertension		
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;		
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;		
KW	cholesterol ester storage; systemic lupus erythematosus; infection;		
KW	severe combined immunodeficiency; malaria; autoimmune disorder; aschma;		
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;		
KW	bone damage; cartilage damage; antiinflammatory disease; coagulation;		
KW	thrombosis; contraceptive; ss.		
OS	homo sapiens.		
XX			
PN	W0200058473-A2.		
XX			
PD	05-OCT-2000.		
XX			
PF	31-MAR-2000; 2000MO-US08621.		
XX			
PR	31-MAR-1999; 99US-0127607.		
PR	02-APR-1999; 99US-0127636.		
PR	05-APR-1999; 99US-0127728.		
PR	30-MAR-2000; 2000US-0540763.		
PA	(CURA-) CURAGEN CORP.		
XX			
PI	Shinkets RA, Leach M;		
XX			
DR	WPI; 2000-602362/57.		
DR	P-PSDB; ABA42783.		
XX			
PT	Novel nucleic acids and peptides derived from open reading frame X,		
PT	useful for treating e.g. cancers, proliferative disorders,		
PT	neurodegenerative disorders and cardiovascular disease -		
XX			
PS	Claim 5; Page 4281-4282; 5507pp; English.		
XX			
CC	AACT74446 to AACT77606 encode the proteins given in AAR40237 to AAR43367		

Db 662 TGACTGTGATCAGCCCTCGAGCTTATCGGACAAAGCTGCAGGCTCAGCATGTGTGT 721
 QY 680 ACCGGAGCTGGGTGCTGTGTTCGAATCGAGTGGCTCAGGCTGCTGGCTCACTGT 739
 Db 722 ACCGGAGCTGGGTGCTGTGTTCGAATCGAGTGGCTCAGGCTGCTGGCTCACTGT 781
 QY 740 GGCTGGCTGGGCCCCACCTGCTTCGAGATGTGCCCTTCTC-----782
 Db 782 GGCTGGCTGGGCCCCACCTGCTTCGAGATGTGCCCTTCTCAGTGCATCCCCACCCC 841
 QY 783 -AGCCCTGTACTTACCTTTGACCTGTGAAGCTGGCTCAATGGTTCAGGCCGA 841
 Db 842 AAGCCCTGTACTTACCTTTGACCTGTGAAGCTGGCTCAATGGTTCAGGCCGA 901
 QY 842 AGGACACAGCTTCTGTGGGCTAGCTTTGTGGCTGGTGGCTATCTCAGGACGGTGGCTG 901
 Db 902 AGGACACAGCTTCTGTGGGCTAGCTTTGTGGCTGGTGGCTATCTCAGGACGGTGGCTG 961
 QY 902 CAGTGTGACTTACCTTTGACCTGTGAAGCTGGCTCAATGGTTCAGGCCGA 961
 Db 962 CAGTGTGACTTACCTTTGACCTGTGAAGCTGGCTCAATGGTTCAGGCCGA 1021
 QY 962 TGAGGCTGTGAGAGTGAACCCCTCGATGTGACTTCCACCTGGCTGGCTGGCTGGAGGA 1021
 Db 1022 TGAGGCTGTGAGAGTGAACCCCTCGATGTGACTTCCACCTGGCTGGCTGGAGGA 1081
 QY 1022 TCCGGGCGAGTGGGGACCAAGGACTTTTTCAGGCTTCTTCCGATCATCAAG 1081
 Db 1082 TCCGGGCGAGTGGGGACCAAGGACTTTTTCAGGCTTCTTCCGATCATCAAG 1141
 QY 1082 CTGCCCTCTCTGTGCCATCATGATCAGCAGTATGAGTTCGGGAAAGCTTCTCCGA 1141
 Db 1142 CTGCCCTCTCTGTGCCATCATGATCAGCAGTATGAGTTCGGGAAAGCTTCTCCGA 1201
 QY 1142 GGCTGAACAGAGCCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGACCCCTCT 1201
 Db 1202 GGCTGAACAGAGCCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGACCCCTCT 1261
 QY 1202 CTCCACGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1261
 Db 1262 CTCCACGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1321
 QY 1262 AGGAGGGGGCTTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1321
 Db 1322 AGGAGGGGGCTTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1381
 QY 1322 TGGGGGCGCCAGCAGCTTCTCAGACACAACTTCTTCTGCTGCTCAGTCTGGGGATCA 1381
 Db 1382 TGGGGGCGCCAGCAGCTTCTCAGACACAACTTCTTCTGCTGCTCAGTCTGGGGATCA 1441
 QY 1382 TCACCTTACCCACCCCAAGTTCAGACCAATCTTCCAGCTGCCCTTCTGTTTCCC 1441
 Db 1442 TCACCTTACCCACCCCAAGTTCAGACCAATCTTCCAGCTGCCCTTCTGTTTCCC 1501
 QY 1442 TGTGTTGTGTAGCTGGGCTGTCTCCAGGACCAAGAGGCTCAGCCTGGTGTAGTC 1501
 Db 1502 TGTGTTGTGTAGCTGGGCTGTCTCCAGGACCAAGAGGCTCAGCCTGGTGTAGTC 1561
 QY 1502 TCCCTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1546
 Db 1562 TCCCTGACCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1606

RESULT 4

AAI71019

ID AAI71019 standard; cDNA; 1114 BP.

XX AAI71019;

AC AAI71019;

DT 18-MAR-2002 (first entry)

XX Human mitochondrial carrier protein CGI-69L cDNA.

DE Human mitochondrial carrier protein CGI-69L cDNA.

XX Human mitochondrial carrier protein CGI-69L cDNA.

KW CGI-69L; mitochondrial carrier protein; human; metabolic disease;
 KW obesity; cachexia; tumor; cancer; infection; immunomodulator;
 KW antitumor; viricide; antibacterial; anorectic; antidiabetic;
 KW brown adipose tissue; diagnosis; gene therapy; splice variant; ss.
 XX Homo sapiens.
 XX OS
 XX Key Location/Qualifiers
 FT CDS 14..1093 /*tag= a
 FT WO200198355-A2.
 PD 27-DEC-2001.
 PF 22-JUN-2001; 2001WO-US20117.
 XX 22-JUN-2000; 2000US-213307P.
 PR (GETH) GENENTECH INC.
 PA (CURA-) CURAGEN CORP.
 XX Lewin D, Adams SH, Yu XX;
 PI WPI; 2002-114569/15.
 DR P-PSDB; AAM50569.
 XX

CGI-69 polypeptides and polynucleotides useful for treating metabolic disorders including cachexia, obesity, diabetes and cancers -
 Claim 1(a); Page 9-10; 107pp; English.

The present sequence is that of cDNA encoding human CGI-69L (see AAM50569), a novel splice variant of human CGI-69. Various CGI-69 clones were isolated from human liver upon PCR amplification and cloning. Numerous clones diverged from the previously known CGI-69 sequence in that they encoded an 8-amino acid insert preceded by a W44L change, and were termed the 'long version' isoform or CGI-69L. The invention relates to a novel characterization of CGI-69 as a mitochondrial carrier protein, the discovery of CGI-69 as a discovery of the alteration of the mitochondrial membrane potential through overexpression of carboxy-FLAG-tagged CGI-69. CGI-69 (including splice variant) nucleic acids and proteins are useful for diagnosing and treating metabolic diseases in humans, including obesity, cachexia and diabetes. Evidence for an important function for CGI-69 in modifying mitochondrial membrane potential in brown adipose tissue (BAT) is presented. The mouse orthologue of CGI-69 is up-regulated in cold-treated BAT. CGI-69 may be involved in cellular thermogenic uncoupling and, therefore, may be used to diagnose and treat specific perturbations in metabolic pathways. Altering the expression of CGI-69 through gene therapy provides a means of treating metabolic diseases, such as obesity or cachexia, or of increasing or decreasing body weight. Decreasing the activity of CGI-69 can be used to treat cachexia, tumors, cancers, viral infections and bacterial infections, while increasing its activity can be used to treat obesity, tumors, cancers, viral infections and bacterial infections (all claimed). Methods are also claimed for determining whether a compound up-regulates or down-regulates expression of a CGI-69 gene, and of screening for a mutation in the CGI-69 gene.

Sequence 1114 BP; 194 A; 349 C; 332 G; 239 T; 0 other;

Query Match 68.2%; Score 1054.4; DB 24; Length 1114;
 Best Local Similarity 97.8%; Pred. No. 8.9e-250;
 Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 105 CTGAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGGT 164
 Db 1 CTGAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGGT 60
 QY 165 GGCTCAGGACCGGGCTGTGGTACCTCTCTTTCATGACCCCTGGACCTGGTGA 224

Db 61 GGCCTCAGGCACCGGGCTGTGGTACCTCTCTTCATGACACCCCTGGAGCTGGTGA 120
 QY 225 GGTTCGCTCAGTCTCAGCGGCTCCATGCGCAGCGAGCTGATGCTTCTCCAGACT 284
 Db 121 GGTTCGCTCAGTCTCAGCGGCTCCATGCGCAGCGAGCTGATGCTTCTCCAGACT 180
 QY 285 GTGAGCTCTCTCTATACCAAAAT-----GGAGTGGCTCTCT 320
 Db 181 GTGAGCTCTCTCTATACCAAAATGGCCCTCTCTCCAAATCCACAGGGAAGTGCCTCT 240
 QY 321 GTATTGCAATGGTCTCTGAGCCTCTGTACTGTGCCCAAAATGGTGGCCGCTGTGCCAC 380
 Db 241 GTATTGCAATGGTCTCTGAGCCTCTGTACTGTGCCCAAAATGGTGGCCGCTGTGCCAC 300
 QY 381 CTGTTTCAAGACCTTACCGCTTCTCACTGGCACCATGATGCTCTGTGGAAGATCTGTAG 440
 Db 301 CTGTTTCAAGACCTTACCGCTTCTCACTGGCACCATGATGCTCTGTGGAAGATCTGTAG 360
 QY 441 GCACGAGGCACACGAGCCTCTGAGCGGCTCTCCCGCACCCCTGGTGTGATGATCTGTGC 500
 Db 361 GCACGAGGCACACGAGCCTCTGAGCGGCTCTCCCGCACCCCTGGTGTGATGATCTGTGC 420
 QY 501 AGCTACGCGCATCTACTGCTGCTATGACCACTGAAGGCTCTCTGTGTGTGTGAGC 560
 Db 421 AGCTACGCGCATCTACTGCTGCTATGACCACTGAAGGCTCTCTGTGTGTGTGAGC 480
 QY 561 CTTGACCTCTGACCTTACGACCATGCTGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCT 620
 Db 481 CTTGACCTCTGACCTTACGACCATGCTGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCT 540
 QY 621 GACTGTGATCAGCCCTCTGAGCTTATGCGGCAAAAGCTGACGCTCAGCATGTGCTA 680
 Db 541 GACTGTGATCAGCCCTCTGAGCTTATGCGGCAAAAGCTGACGCTCAGCATGTGCTA 600
 QY 681 CCGGAGCTGGTGGCTCTGTTCGAATGCACTGAGTGGCTGAGGTGGCTGGCGCTCACTGTG 740
 Db 601 CCGGAGCTGGTGGCTCTGTTCGAATGCACTGAGTGGCTGAGGTGGCTGGCGCTCACTGTG 660
 QY 741 GCTGGCTGGGCGCCCACTGCGCTTTCGAGATGTCCTTCTCAGCGCTGACTGCTCAA 800
 Db 661 GCTGGCTGGGCGCCCACTGCGCTTTCGAGATGTCCTTCTCAGCGCTGACTGCTCAA 720
 QY 801 CTATGAGCTGTGAAGAGCTGGCTCAATATGGTTTCAAGCGGAGGACAGACTTCTGTGG 860
 Db 721 CTATGAGCTGTGAAGAGCTGGCTCAATATGGTTTCAAGCGGAGGACAGACTTCTGTGG 780
 QY 861 CATGAGCTTGTGGCTGGTGGCTCTCAGGAGCTGGCTGAGTGGCTGACTTACCTTT 920
 Db 781 CATGAGCTTGTGGCTGGTGGCTCTCAGGAGCTGGCTGAGTGGCTGACTTACCTTT 840
 QY 921 TCACGTGTGAAGACCCCAAGCCAGGTGCTCTGGGAGCGATGAGGCTGTGAGAGTGAA 980
 Db 841 TCACGTGTGAAGACCCCAAGCCAGGTGCTCTGGGAGCGATGAGGCTGTGAGAGTGAA 900
 QY 981 CCCCCGTGATGAGCTCACTGCTGCTGCTGCTGCGAGGATCGGGCGGAGTGGGCGAC 1040
 Db 901 CCCCCGTGATGAGCTCACTGCTGCTGCTGCTGCGAGGATCGGGCGGAGTGGGCGAC 960
 QY 1041 CAAGGAGCTTTTTCAGGCTCTCTCTCGGATCATCAGGCTGCCCTCTCTGTGCCAT 1100
 Db 961 CAAGGAGCTTTTTCAGGCTCTCTCTCGGATCATCAGGCTGCCCTCTCTGTGCCAT 1020
 QY 1101 CATGATCAGCACTTATGATTTGGGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCT 1160
 Db 1021 CATGATCAGCACTTATGATTTGGGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCT 1080
 QY 1161 TCTGGCGGCTGAAGGGGCAAGGAGGAC 1194
 Db 1081 TCTGGCGGCTGAAGGGGCAAGGAGGAC 1114

RESULT 5
 AAA06606

ID AAA06606 standard; cDNA; 537 BP.
 XX AAA06606;
 XX 13-JUN-2000 (first entry)
 XX Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:387.
 XX Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
 KW immunogenic; cytostatic; vaccine; ss.
 XX Homo sapiens.
 XX WO200004149-A2.
 XX 27-JAN-2000.
 XX 14-JUL-1999; 99WO-US15838.
 XX 14-JUL-1998; 98US-0115453.
 PR 14-JUL-1998; 98US-0116134.
 PR 23-SEP-1998; 98US-0159812.
 PR 23-SEP-1998; 98US-0159822.
 PR 15-JAN-1999; 98US-0232149.
 PR 15-JAN-1999; 98US-0232880.
 PR 09-APR-1999; 99US-0288946.
 XX (CORI-) CORIXA CORP.
 PA Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;
 XX WPI; 2000-171268/15.
 XX New polypeptide useful for treating and diagnosing prostate cancer
 PT comprises an immunogenic portion of prostate tumor protein -
 XX Claim 50; Page 235-236; 263pp; English.
 CC The present invention describes isolated polypeptides, comprising an
 CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
 CC and polynucleotides encoding them have cytostatic activity and can be
 CC used in vaccines and in gene therapy. The polypeptides and
 CC polynucleotides encoding them, antigen presenting cells which express
 CC the polypeptides, antibodies against the polypeptides and v cines
 CC comprising them can be used for inhibiting the development of prostate
 CC cancer in a patient. The polypeptides can be used to generate antibodies
 CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
 CC the polynucleotides encoding the polypeptides can be used as a probe or
 CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
 CC AAY82000 to AAY82020 represent sequences used in the exemplification of
 CC the present invention.

Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 33.8%; Score 522; DB 21; Length 537;
 Best Local Similarity 100.0%; Pred. No. 8.7e-119;
 Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTCGGGACCAAGGAGCTTTTGGAGGCTTCTTCGAGGCTTCTTCGAGTATCAGAGCTG 1084
 Db 1 GGGCCGAGTCGGGACCAAGGAGCTTTTGGAGGCTTCTTCGAGGCTTCTTCGAGTATCAGAGCTG 60
 QY 1085 CCCCCCTCTGTGCCATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTCAGAGGC 1144
 Db 61 CCCCCCTCTGTGCCATCATGATCAGCACCTATGAGTTCGGCAAAAGCTTCTTCAGAGGC 120
 QY 1145 TGAACCCAGGACCGGCTTCTGGGCGGCTGAAGGGGCAAGGAGGACCCCGTCTCTC 1204
 Db 121 TGAACCCAGGACCGGCTTCTGGGCGGCTGAAGGGGCAAGGAGGACCCCGTCTCTC 180
 QY 1205 CCACGGATGGGAGAGGGCAGGAGGAGACCCAGGCCAGTGCCTTTCTCAGACTGAGG 1264
 Db 181 CCACGGATGGGAGAGGGCAGGAGGAGACCCAGGCCAGTGCCTTTCTCAGACTGAGG 240

QY 1265 GAGGGGCTTGTTCCTTCCCTCCCGGCGAAGCTCCAGGCGAGGGCTGTCCCTCTGG 1324
 DB 241 GAGGGGCTTGTTCCTTCCCTCCCGGCGAAGCTCCAGGCGAGGGCTGTCCCTCTGG 300
 QY 1325 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATATCA 1384
 DB 301 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATATCA 360
 QY 1385 CTTACCCACCCCAAGTTCAAGACCAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 1444
 DB 361 CTTACCCACCCCAAGTTCAAGACCAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 420
 QY 1445 GTTTCCTGTAGCTGGGCATGCTCCAGAACCAAGAGCCCTCAAGCTGGTGTAGTCTCC 1504
 DB 421 GTTTCCTGTAGCTGGGCATGCTCCAGAACCAAGAGCCCTCAAGCTGGTGTAGTCTCC 480
 QY 1505 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1546
 DB 181 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 522

RESULT 6
 AAS63815
 ID AAS63815 standard; cDNA; 537 BP.
 AC AAS63815;
 XX
 DT 29-JAN-2002 (first entry)
 DE Human prostate cDNA sequence #357.
 KW Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
 OS Homo sapiens.
 XX
 PW WO200173032-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-US09919.
 XX
 PR 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PF 10-AUG-2000; 2000US-0636215.
 P 19-AUG-2000; 2000US-0651236.
 PK 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX
 (CORI-) CORIXA CORP.
 XX
 XU Xu J, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;
 PI Li SX, Wang A, Skeiky IAW, Hepler WT, Henderson RA;
 XX
 DR WPI; 2001-639232/73.
 XX

PT New human prostate-specific polypeptides and polynucleotides useful for
 PT the diagnosis and treatment of cancer, especially prostate cancer -
 XX
 PS Claim 1; Page 361; 579pp; English.
 XX
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding

CC T cells specific for a tumour protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polynucleotide of the invention.
 XX

QY Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 33.8%; Score 522; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 8.7e-119;
 Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTGGGACCAAGGACATCTTTCAGGCTTCTTCTCGATCATCAAGGCTG 1084
 DB 1 GGGCCGAGTGGGACCAAGGACATCTTTCAGGCTTCTTCTCGATCATCAAGGCTG 60
 QY 1085 CCCCCTCTGTGCCATCATGATCAGACCTATGATTTGSCAAAGCTTCTCCAGAGGC 1144
 DB 61 CCCCCTCTGTGCCATCATGATCAGACCTATGATTTGSCAAAGCTTCTCCAGAGGC 120
 QY 1145 TGAACACAGGACCGGCTTCTTGGGCGCTGAAAGGGGCAAGGAGGACCCCGTCTCTC 1204
 DB 121 TGAACACAGGACCGGCTTCTTGGGCGCTGAAAGGGGCAAGGAGGACCCCGTCTCTC 180
 QY 1205 CCACGATGGGAGAGGCGCAGGAGACCCAGCAAGTGCCTTTCTCAGCACTGAGG 1264
 DB 181 CCACGATGGGAGAGGCGCAGGAGACCCAGCAAGTGCCTTTCTCAGCACTGAGG 240
 QY 1265 GAGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGCGAGGCTGTCCCTCTGG 1324
 DB 241 GAGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGCGAGGCTGTCCCTCTGG 300
 QY 1325 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATATCA 1384
 DB 301 GCGGCCAGACACTTCTCAGACACAACTTCTTCTGCTGCTCCAGTCTGGGGATATCA 360
 QY 1385 CTTACCCACCCCAAGTTCAAGACCAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 1444
 DB 361 CTTACCCACCCCAAGTTCAAGACCAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 420
 QY 1445 GTTTCCTGTAGCTGGGCATGCTCCAGAACCAAGAGCTTAAAGATGATGAATTC 1546
 DB 421 GTTTCCTGTAGCTGGGCATGCTCCAGAACCAAGAGCTTAAAGATGATGAATTC 522

RESULT 7
 AAH93722
 ID AAH93722 standard; cDNA; 537 BP.
 XX
 AC AAH93722;
 XX
 DT 04-OCT-2001 (first entry)
 XX
 DE Human prostate-specific cDNA sequence CGI-69.
 XX
 KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
 KW cytostatic; gene therapy; metastasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200151633-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001WO-US01574.
 XX
 PR 14-JAN-2000; 2000US-0483672.
 XX

PA (CORI-) CORIXA CORP.
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Panger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
 PI Wang A, Meagher MJ;
 XX WPI; 2001-425873/45.
 DR
 XX New polynucleotide encoding a prostate-specific protein, for
 PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines -
 XX
 PS Claim 1; Page 359-360; 543pp; English.
 XX
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
 CC AAH01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.
 XX
 XX Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 33.8%; Score 522; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 8.7e-119;
 Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1025 GGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGATCATCAAGGCTG 1084
 DB 1 GGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGATCATCAAGGCTG 60
 QY 1085 CCCCTCTCTGTCATCATGATCAGCACCTATGATTTGGCAAAAGCTTCTTCAGAGGC 1144
 DB 61 CCCCTCTCTGTCATCATGATCAGCACCTATGATTTGGCAAAAGCTTCTTCAGAGGC 120
 QY 1145 TGAACCCAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGCCCGCTCTCTC 1204
 DB 121 TGAACCCAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGCCCGCTCTCTC 180
 QY 1205 CCACGGATGGGAGAGGCGAGGAGAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 1264
 DB 181 CCACGGATGGGAGAGGCGAGGAGAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 240
 QY 1265 GAGGGGGCTTGTTCCTTCCCTCCCGGCAAGCTCCAGGAGGGCTGCTCTCTGG 1324
 DB 241 GAGGGGGCTTGTTCCTTCCCTCCCGGCAAGCTCCAGGAGGGCTGCTCTCTGG 300
 QY 1325 GGCGCCAGCACTTCTTCAGACAACTTCTCTGCTGCTCCAGTGGTGGGATCATCA 1384
 DB 301 GGCGCCAGCACTTCTTCAGACAACTTCTCTGCTGCTCCAGTGGTGGGATCATCA 360
 QY 1385 CTTACCCACCCCAAGTCTCAAGCAAAATCTTCAGTGCCCGCTTCGTGTTCCTGT 1444
 DB 361 CTTACCCACCCCAAGTCTCAAGCAAAATCTTCAGTGCCCGCTTCGTGTTCCTGT 420
 QY 1445 GTTCTCTAGCTGGGATCTCTCCAGGAACCCAGAGCCCTCAGCCTGTGTAGTCTCC 1504
 DB 421 GTTCTCTAGCTGGGATCTCTCCAGGAACCCAGAGCCCTCAGCCTGTGTAGTCTCC 480
 QY 1505 CTGACCTTGTAAATCTTAACTCTTAAAGATGATGAATTC 1546
 DB 481 CTGACCTTGTAAATCTTAACTCTTAAAGATGATGAATTC 522

RESULT 8

AAH85036
 ID AAH85036 standard; cDNA; 537 BP.
 XX
 AC AAH85036;
 XX
 XX 25-SEP-2001 (first entry)
 XX
 DE Human prostate-specific cDNA sequence CGI-69.
 XX
 KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
 KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
 KW prostate specific antigen; PSA; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200134802-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US30904.
 XX
 PF 12-NOV-1999; 99US-0439313.
 PF 18-NOV-1999; 99US-0443686.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
 XX WPI; 2001-308785/32.
 XX
 PT Isolated polypeptide comprising at least an immunogenic portion of a
 PT prostate-specific protein, useful in the diagnosis and therapy of
 PT prostate cancer -
 XX
 PS Claim 31; Page 259; 325pp; English.
 XX
 CC The present invention describes an isolated polypeptide (PI) comprising
 CC at least an immunogenic portion of a prostate-specific protein, or its
 CC variant. Also described are polynucleotides (NI) encoding (PI). (PI) and
 CC (NI) have cytostatic activity and can be used in vaccine production.
 CC The polypeptides, nucleic acids and antibodies from the present
 CC invention are useful in the diagnosis and therapy of prostate cancer.
 CC prostate specific genes P704P, P712P, P774P, P775P and B305D are located
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
 CC region. Prostate specific antigen (PSA) P501S was located on
 CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
 CC polynucleotide and polypeptide sequences used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 33.8%; Score 522; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 8.7e-119;
 Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1025 GGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGATCATCAAGGCTG 1084
 DB 1 GGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGATCATCAAGGCTG 60
 QY 1085 CCCCTCTCTGTCATCATGATCAGCACCTATGATTTGGCAAAAGCTTCTTCAGAGGC 1144
 DB 61 CCCCTCTCTGTCATCATGATCAGCACCTATGATTTGGCAAAAGCTTCTTCAGAGGC 120
 QY 1145 TGAACCCAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGCCCGCTCTCTC 1204
 DB 121 TGAACCCAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGCCCGCTCTCTC 180
 QY 1205 CCACGGATGGGAGAGGCGAGGAGAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 1264
 DB 181 CCACGGATGGGAGAGGCGAGGAGAGCCAGCAAGTGCCTTTCTCAGCACTGAGG 240
 QY 1265 GAGGGGGCTTGTTCCTTCCCTCCCGGCAAGCTCCAGGAGGGCTGCTCTCTGG 1324


```

Db 241 GAGGGGCTGTTCCCTCCCTCCGGGCAAGCTCCAGGGCAGGGCTGTCCTCTCTGG 300
Qy 1325 GCGGCCAGCAGCTCTCCAGACAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1384
Db 301 GCGGCCAGCAGCTCTCCAGACAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Qy 1385 CTTACCCACCCCGGAGTTCAGACCAAACTTCTCTCTCTCTCTCTCTCTCTCTCT 1444
Db 361 CTTACCCACCCCGGAGTTCAGACCAAACTTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 1445 GTTTCGTAGTGGGATGCTCCAGGAACCAAGAGCCCTCAGCCCTGCTGTAGTCTCC 1504
Db 421 GTTTCGTAGTGGGATGCTCCAGGAACCAAGAGCCCTCAGCCCTGCTGTAGTCTCC 480
Qy 1505 CTGACCCCTGTTAACTTCTTAAGTCTAAAGATGATGAATTC 1546
Db 481 CTGACCCCTGTTAACTTCTTAAGTCTAAAGATGATGAATTC 522

RE
Abh.../87
ID AA02787 standard; cDNA; 537 BP.
XX
AC AA02787;
XX
DT 14-JUN-2001 (first entry)
XX
DE Prostate tumour antigen cDNA sequence for CGI-69.
XX
KW Human; prostate tumour antigen; prostate cancer; therapy; diagnosis;
KW prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200125272-A2.
XX
PD 12-APR-2001.
XX
PF 04-OCT-2000; 2000WO-US27464.
XX
PR 04-OCT-1999; 99US-0157455.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Skeiky YAW, Reed SG, Cheever MA;
XX
WP1; 2001-245062/25.
XX
X Prostate specific protein and its encoding polynucleotide, useful for
PT the treatment and diagnosis of prostate cancer -
XX
PS Claim 50; Page 244; 276pp; English.
XX
CC The present invention describes an isolated polypeptide (I) comprising
CC at least an immunogenic portion of a prostate tumour antigen protein or
CC its variant. (I) have cytostatic activity and can be used in vaccine
CC production. (I), prostate tumour antigen polynucleotides, an antigen
CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC pharmaceutical composition containing (I) are useful for inhibiting the
CC development of cancer in a patient. Antibodies specific for prostate
CC specific proteins and oligonucleotides that hybridise to a
CC polynucleotide that encodes a prostate specific protein are useful
CC for detecting the presence or absence of a cancer or monitoring the
CC progression the progression of a cancer, especially prostate cancer.
CC AA02422 to AAB7872, AAB7498 to AAB74821 and AAB74830 are sequences
CC used in the exemplification of the present invention.
XX
SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 33.8%; Score 522; DB 22; Length 537;
Best Local Similarity 100.0%; Pred. No. 8.7e-119;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1025 GGGCCGAGTCGGGCAACAGGGACTCTTTGCAAGGCTTCTTCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCGAGTCGGGCAACAGGGACTCTTTGCAAGGCTTCTTCTCGGATCATCAAGGCTG 60
Qy 1085 CCCCCTCTCTGCCATCATGATCAGCACCTATGAGTTCGCAAAAGCTTCTTCCAGAGGC 1144
Db 61 CCCCCTCTCTGCCATCATGATCAGCACCTATGAGTTCGCAAAAGCTTCTTCCAGAGGC 120
Qy 1145 TGAACACAGACCGGCTTCTGCGCGGCTGAAAGGGCAAGAGGAGGAGGAGGAGGAGG 1204
Db 121 TGAACACAGACCGGCTTCTGCGCGGCTGAAAGGGCAAGAGGAGGAGGAGGAGGAGG 180
Qy 1205 CCACGATCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1264
Db 181 CCACGATCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Qy 1265 GAGGGGGCTTGTTCCTTCCCTCCCGGGCAACAGCTCCAGGCGAGGGCTGTCCCTCTGG 1324
Db 241 GAGGGGGCTTGTTCCTTCCCTCCCGGGCAACAGCTCCAGGCGAGGGCTGTCCCTCTGG 300
Qy 1325 GCGGCCAGCAGCTCTCCAGACAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1384
Db 301 GCGGCCAGCAGCTCTCCAGACAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Qy 1385 CTTACCCACCCCGGAGTTCAGACCAAACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1444
Db 361 CTTACCCACCCCGGAGTTCAGACCAAACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 1445 GTTTCGTAGTGGGATGCTCCAGGAACCAAGAGCCCTCAGCCCTGCTGTAGTCTCC 1504
Db 421 GTTTCGTAGTGGGATGCTCCAGGAACCAAGAGCCCTCAGCCCTGCTGTAGTCTCC 480
Qy 1505 CTGACCCCTGTTAACTTCTTAAGTCTAAAGATGATGAATTC 1546
Db 481 CTGACCCCTGTTAACTTCTTAAGTCTAAAGATGATGAATTC 522

RESULT 10
ABL95186
ID ABL95186 standard; cDNA; 537 BP.
XX
AC ABL95186;
XX
DT 19-JUL-2002 (first entry)
XX
DE Human CGI-69 cDNA sequence SEQ ID NO 387.
XX
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN US2002022248-A1.
XX
PD 21-FEB-2002.
XX
12-JAN-2001; 2001US-0759143.
XX
25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
PR 09-FEB-1998; 98US-0020956.
PR 25-FEB-1998; 98US-0030607.
PR 14-JUL-1998; 98US-0115453.
PR 23-SEP-1998; 98US-0159812.
PR 15-JAN-1999; 99US-0232149.
PR 03-APR-1999; 99US-0288946.
PR 13-JUL-1999; 99US-0352616.
PR 18-NOV-1999; 99US-0439313.
PR 18-NOV-1999; 99US-0443686.
PR 14-JAN-2000; 2000US-0483672.
PR 27-MAR-2000; 2000US-0536857.
PR 09-MAY-2000; 2000US-0568100.

```


PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX (XUJ)/ XU J.
 PA (DILL)/ DILLON D C.
 PA (MITC)/ MITCHAM J L.
 PA (HARL)/ HARLOCKER S L.
 PA (JIAN)/ JIANG Y.
 PA (KALO)/ KALOS M D.
 PA (FANG)/ FANGER G R.
 PA (RETT)/ RETTER M W.
 PA (STOL)/ STOLK J A.
 PA (DAYC)/ DAY C H.
 PA (VEDV)/ VEDVICK T S.
 PA (CART)/ CARTER D.
 PA (LISX)/ LI S X.
 PA (WANG)/ WANG A.
 PA (SKEL)/ SKELLY Y A W.
 PA (HEPL)/ HEPLER W T.
 PA (HEND)/ HENDERSON R A.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX WPI; 2002-255649/30.
 DR
 XX New prostate-specific polynucleotides for diagnosing and treating
 PT diseases, in particular prostate cancer, and as markers for the
 PT progression of cancer
 XX
 PS Claim 1; SEQ ID NO 387; 87pp; English.
 XX
 CC The present invention provides prostate-specific coding sequences and
 CC their encoded proteins. These can be used in the diagnosis and treatment
 CC of cancers, particularly prostate cancer. The present sequence is a cDNA
 CC described in the invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 33.8%; Score 522; DB 24; Length 537;
 Best Local Similarity 100.0%; Pred. No. 8.7e-119;
 Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1025 GGGCCGAGTCGGGACCAAGGACTCTTTGACAGGCTCTCTCGATCATCAAGGCTG 1084
 Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTGACAGGCTCTCTCGATCATCAAGGCTG 60
 QY 1085 CCCCTCTGCGCATCATGATCAGCACTATGATTCGGCAAGCTCTTCAGAGGC 1144
 Db 61 CCCCTCTGCGCATCATGATCAGCACTATGATTCGGCAAGCTCTTCAGAGGC 120
 QY 1145 TGACAGGACCGCTCTGGGGCTGTAAGGGGCAAGGAGGACCGGCTCTC 1204
 Db 121 TGACAGGACCGCTCTGGGGCTGTAAGGGGCAAGGAGGACCGGCTCTC 180
 QY 1205 CCACGGATGGGAGGAGGAGGAGGACCCAGCAAGTGCCTTTCTCAGACTGAGG 1264
 Db 181 CCACGGATGGGAGGAGGAGGAGGACCCAGCAAGTGCCTTTCTCAGACTGAGG 240
 QY 1265 GAGGGGGCTGTTTCCCTTCCCTCCGGGCAAGAGCTCCAGGAGGCTGTCCTCTGG 1324
 Db 241 GAGGGGGCTGTTTCCCTTCCCTCCGGGCAAGAGCTCCAGGAGGCTGTCCTCTGG 300
 QY 1325 GGGGGCCAGCACTTCTCAGACACAACTTCTCTGCTGCTCCAGTCGTGGGATCATCA 1384
 Db 301 GGGGGCCAGCACTTCTCAGACACAACTTCTCTGCTGCTCCAGTCGTGGGATCATCA 360

QY 1385 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTGCCCCCTTCGTGTTCCCTGT 1444
 Db 361 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTGCCCCCTTCGTGTTCCCTGT 420
 QY 1445 GTTTGCTGTAGCTGGGCATGTCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 1504
 Db 421 GTTTGCTGTAGCTGGGCATGTCTCCAGGAACCAAGAGCCCTCAGCCTGGTGTAGTCTCC 480
 QY 1505 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1546
 Db 481 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 522
 RESULT 11
 ABK34735
 ID ABK34735 standard; cDNA; 2170 BP.
 XX
 AC ABK34735;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA for novel secreted protein, SEQ ID 504.
 XX
 KW Human; ss; gene; secreted protein; immune deficiency; viral infection;
 KW bacterial infection; fungal infection; autoimmune disorder; burn;
 KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
 KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
 KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
 KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
 KW lymphoid cell deficiency.
 XX
 OS Homo sapiens.
 XX
 PN WO200177290-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US10295.
 XX
 PR 06-APR-2000; 2000US-194941P.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;
 XX WPI; 2002-179323/23.
 XX
 PT Six hundred and twenty five polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for
 PT treating immune deficiencies and disorders such as autoimmune disorders
 PT
 PS Claim 1; Page 269-270; 339pp; English.
 XX
 CC The invention relates to 625 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins, their complements and sequences that hybridise to them.
 CC Also included are a vector comprising the polynucleotide, a host cell
 CC transformed with the vector, the proteins encoded by the
 CC polynucleotides, antibodies that bind to the proteins and identification
 CC of modulators of the proteins or the expression of the polynucleotide.
 CC The polynucleotides can be used as probes for the identification
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 CC and proteins can also be used as nutritional supplements. The protein
 CC is useful in the treatment of various immune deficiencies and disorders
 CC such as viral infections, bacterial infections, fungal infections,
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 CC autoimmune thyroiditis and diabetes) and allergic reactions and
 CC conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's

CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment
 CC of burns, incisions and ulcers. The proteins are also useful for
 CC regulating haematopoiesis, for treating myeloid or lymphoid cell
 CC deficiencies. The present sequence is one of the 625 cDNA sequences
 CC encoding a secreted protein.
 XX

SQ Sequence 2170 BP; 407 A; 663 C; 636 G; 460 T; 4 other;

Query Match		33.5%	Score 517.2;	DB 24;	Length 2170;
Best Local Similarity		64.0%	Pred. No. 2e-117;		
Matches 1290;		Conservative	0;	Mismatches	8;
				Indels	718;
				Gaps	5;
QY	100	CAGGACTGAGCTTCAAGATGGCTGACAGGACCTTGGGGCATCAGCCCTCCAGCAA	159		
DB	148	CAGGACTGAGCTTCAAGATGGCTGACAGGACCTTGGGGCATCAGCCCTCCAGCAA	207		
QY	160	ATGTTGGCTCAGGACCGGGCTGGTTTACTCTCTCTTCATGACACCCCTGGAGTG	219		
DI	208	ATGTTGGCTCAGGACCGGGCTGGTTTACTCTCTCTTCATGACACCCCTGGAGTG	267		
QY	220	GTGAAGTTTGGCTCAGGCTCAGGCTTCCATGGCCAGGAGCTGATGCTTCTCC	279		
DB	268	GTGAAGTTTGGCTCAGGCTCAGGCTTCCATGGCCAGGAGCTGATGCTTCTCC	327		
QY	280	AGACTTGGAGCTTCTCTATACCAATGGAAGTCTCTGTATTCGAATGTGTCTG	339		
DB	328	AGACTTGGAGCTTCTCTATACCAATGGAAGTCTCTGTATTCGAATGTGTCTG	387		
QY	340	GAGCTCTGTACTGTGCTCCCAATGTTGGCCCTGTGCACTGTGTTTCAAGACCTTCC	399		
DB	388	GAGCTCTGTACTGTGCTCCCAATGTTGGCCCTGTGCACTGTGTTTCAAGACCTTCC	447		
QY	400	CGTTTCACTGGCACAATGATGCTTCTGTGAAGTCTGTGAGCACTGAGGACCAAGACC	459		
DB	448	CGTTTCACTGGCACAATGATGCTTCTGTGAAGTCTGTGAGCACTGAGGACCAAGACC	507		
QY	460	CTCTGAGCGGCTCCCGCCACCTGTGATGACTGTGCCAGCTACCGCCATCTACTTC	519		
DB	508	CTCTGAGCGGCTCCCGCCACCTGTGATGACTGTGCCAGCTACCGCCATCTACTTC	567		
QY	520	ACTGCTATGACCAACTGAAGCTTCTCTGTGTTGAGGCTTGACTCTGACCTCTAC	579		
DB	568	ACTGCTATGACCAACTGAAGCTTCTCTGTGTTGAGGCTTGACTCTGACCTCTAC	627		
QY	580	GCACCATGTGCTGGCGCTGSCCGCC	610		
D	628	GCACCATGTGCTGGCGCTGSCCGCC	687		
QY	611	-----	610		
DB	688	CTCTGTCCCCCACTCGGCCCAACCCAGTCAGTCCAGCTCCAGCTCCAGCTCCGCA	747		
QY	611	-----	610		
DB	748	CCCTAGGATAGGGGGTGTCTAGTGTCCCAAGACTGGGACAGAGGTAGGTAGTGA	807		
QY	611	-----TGGGCA	616		
DB	808	TGTAAGACCTGTGATCTCACTTCTCCCAATTGTACCCCAATCCCCCACACAGTGGCA	867		
QY	617	CCGTGATGTGATACAGCCCTGGAGTTATGGGACAAAGTGGAGCTCAGCATGTGT	676		
DB	868	CCGTGATGTGATACAGCCCTGGAGTTATGGGACAAAGTGGAGCTCAGCATGTGT	927		
QY	677	CGTACCGGAGCTGGGTGCTGTGTTCGAATGCACTGAGTGGCTCAGGCTGCGCTCAC	736		
DB	928	CGTACCGGAGCTGGGTGCTGTGTTCGAATGCACTGAGTGGCTCAGGCTGCGCTCAC	987		
QY	737	TGTGGCTGGGCTGGGGCCCACTGCCCTTCGAGATGTGCCCTTCTC	782		
DB	988	TGTGGCTGGGCTGGGGCCCACTGCCCTTCGAGATGTGCCCTTCTCAGGTAGGACCCAA	1047		

QY	783	-----	782
DB	1048	GCTGCAAGGGTGGGTGAGGAAACCTTTGGGTGATGACAAAGGGTCTACAGCCNANGTC	1107
QY	783	-----	782
DB	1108	CAANTTTGTGACTAGTGACCATGAAGTTACTTAGCTTTTCTTGACCCATGGGTTTCTCAT	1167
QY	783	-----	782
DB	1168	CTGTGATTTGGGCTTAGCAGGTAGAGTATTTGGCCAGNCGGTACAGAGATTAAATTCCTAC	1227
QY	783	-----	782
DB	1228	AGAGTGCCTGCTTCTGAGCTGCTATGTAGCAGACAAGACTCACAGCTGGGAGGTGGT	1287
QY	783	-----	782
DB	1288	GGCAGTGAAGGTGGACGGTCTGGGGCCACTTCTTGCCCTTCCCCTCTCTAATCAGTGC	1347
QY	783	-----AGCCCTGTACTGTGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATG	829
DB	1348	ATCCCCCAACCCCAAGCCCTGTACTGTTCACATATGAGCTGGTGAAGAGCTGGCTCAATG	1407
QY	830	GGTTGAGCTCGAAGAGACAGACTTCTCTGTGGCATGAGCTTTGTGGCTGGTGGCATCTCAG	889
DB	1408	GGTTGAGCTCGAAGAGACAGACTTCTCTGTGGCATGAGCTTTGTGGCTGGTGGCATCTCAG	1467
QY	890	GGAC-----	893
DB	1468	GGACGGTGAAGTGGACTGGGCTGAAGTGAAGCCAGCGAGGTCTCCCTCGGGGGGGCTG	1527
QY	894	-----	893
DB	1528	AGGATTTTGGAGATGCCAGAGCAGACCCCATGAGGGCTTTGTCTCCCTCCAGCCGCAA	1587
QY	894	-----	893
DB	1588	GAGCCATCTGGAGTAGGGTGGGGTCCAGGAAGTGGAGCTGAGCCCTCTCCCGGCC	1647
QY	894	-----GGTGGCTGCAAGTCTGACTCTACCTTTGAGCTGGTAAAGACCAACGCCA	944
DB	1648	ACCCACACAGGTGGCTGCAAGTCTGACTCTTACCTTTGAGCTGGTAAAGACCAACGCCA	1707
QY	945	GGTGGCTCTGGAGCATGGAGCTGTGAGAGTGAACCCCTGCACTGGAGTCCACCTG	1004
DB	1708	GGTGGCTCTGGAGCATGGAGCTGTGAGAGTGAACCCCTGCACTGGAGTCCACCTG	1767
QY	1005	GCTGTGCTCGAGGATCCGGCCGAGTCCGGSCACCAAGGGACTCTTTGAGGCTTCTCT	1064
DB	1768	GCTGTGCTCGAGGATCCGGCCGAGTCCGGSCACCAAGGGACTCTTTGAGGCTTCTCT	1827
QY	1065	TCCTCGGATCATCAAGCTGCCCCCTCTCTGTGCACTGATGATGAGTTCGG	1124
DB	1828	TCCTCGGATCATCAAGCTGCCCCCTCTCTGTGCACTGATGATGAGTTCGG	1887
QY	1125	CAAAAGCTTCTTCCAGAGCTGAACCAAGGACCGGCTTCTGGGGCTGAAAGGGGCAAGG	1184
DB	1888	CAAAAGCTTCTTCCAGAGCTGAACCAAGGACCGGCTTCTGGGGCTGAAAGGGGCAAGG	1947
QY	1185	AGGCAAGG--ACCCCGTCTCTCCACCGATGGGAGA--GGGCAAGGAGAGCCAGGCCAA	1241
DB	1948	AGGCAAGGACACCCCTCTCATCCCAAGATGGGAGATGGGAGGAGAGCCAGGCCAA	2007
QY	1242	GTGCTTTTCTCAGCACTGAGGAGGGGGCTTGTTCCTTCCCTCCCGGCGCAAGCT	1301
DB	2008	GTGCTTTTCTCAGCACTGAGGAGGGGGCTTGTTCCTTCCCTCCCGGCGCAAGCT	2067
QY	1302	CCAGGCAAGGGCTGTCCCTCTGGGGGCCAGCACTTCTCTCAGACACAACTTCTCTGC	1361
DB	2068	CCAGGCAAGGGCTGTCCCTCTGGGGGCCAGCACTTCTCTCAGACACAACTTCTCTGC	2127

QY 1362 TGCTCAGTGTGGGATCATCACTTACCCACCCC 1397
 Db 2128 TGCTCAGTGTGGGATCATCACTTACCCACACC 2163

RESULT 12

ABI99647
 ID ABI99647 standard; cDNA; 704 BP.
 XX AC ABI99647;
 XX DT 07-MAR-2002 (first entry)
 XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:675.
 XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 XX KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
 XX OS Mus musculus.
 XX PN WO20018188-A2.
 XX PD 22-NOV-2001.
 XX PF 18-MAY-2001; 2001WO-JP04192.
 XX PR 18-MAY-2000; 2000JP-0145977.
 XX PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX DR WPI; 2002-034733/04.
 XX EX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes -
 XX

Claim 2; Page 1679-1680; 2690pp; English.

CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding
 CC the protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.
 XX

SQ Sequence 704 BP; 132 A; 204 C; 209 G; 159 T; 0 other;

Query Match 30.4%; Score 470.4; DB 24; Length 704;
 Best Local Similarity 83.1%; Pred. No. 4.8e-106;
 Matches 560; Conservative 0; Mismatches 111; Indels 3; Gaps 2;

QY 512 TTTACTTACTGCTATGACCACTGAAGGCTTCTGTGTGCTGAGCCCTGACCTCTG 571
 Db 3 TCGATCCATGGCTTACGACCACTCAAGGCTTCTGTGTGCTGAGCTGACCTCTG 62
 QY 572 ACCTTACGACCCATGTGTGCTGGCGCTGGCCCGCTGGGACCGCTGACTGTGATCA 631
 Db 63 ACCTTACGACCCATGTGTGCTGGCGCTGGCCCGCTGGGACCGCTGACTGTGATCA 120
 QY 632 GCCCCTGGAGCTTATCGGACAAAGCTGCGAGCTCAGAGCTGCTGACCGGAGCTGG 691
 Db 121 GCCCCTGGAGCTGCTGGGACCAAGCTGCGAGCTCAGAGCTGCTGATACCGTGGCTGG 180

QY 692 GTGCTGTGTTCGAATGCAGTGGCTCAGGGTGGCTGGCGCTCACTGTGGTGGGTGGG 751
 Db 181 CTTCCCTGTTCAGCTCGGTGACTCAGGGTGGCTGGCGCTCTCTGTGGTGGGTGGG 240
 QY 752 GCCCACTGCCCTTCGAGATGTCCTTCAGCCCTCTACTGTGTTCAATATGAGCTGG 811
 Db 241 GTCCCACTGCTTCGAGATGTCCTTCAGCTCTGACTGTGTTCAATACAGTTGG 300
 QY 812 TGAAGAGCTGGCTCAATGGGTTCAGGCCGAGACCAAGACTTCTGTGGCATGAGTTTG 871
 Db 301 TGAAGAGCTGGCTGAGCGACTGAGACCAAGACTTCTGTGGCATGAGTTTG 360
 QY 872 TGCTGTGTGGCATCTCAGGGACGCTGCTGAGTGTGCTGACTCTACCTTTGAGCTGGTAA 931
 Db 361 TGCGCGGTGG-ATCTCAGGAATGCTGCTGCCACCTTACCTACCTTCGATGTGGTGA 419
 QY 932 AGACCCAGCGCAGTGTGCTGCGAGCGATGAGGCTGTGAGAGTGAACCCCTGCATG 991
 Db 420 AGACACAGCGACAGATGTCACTGGGAGCGGTGAGGCTGTGAGAGTGAACCCCGCAGAG 479
 QY 992 TGGACTTCCACTGTGCTGCTGCGAGGATCCGGGCGAGTGGGACCAAGGACTCT 1051
 Db 480 TGAATCTCCACTGTGCTGCTGCGAGGATCCGGGCGTGAATCTGGCACCAGGACTCT 539
 QY 1052 TTCCAGGCTTCTTCTCCGATCATCAAGGCTGCCCCCTCTCTGTCATCATGATCAGCA 1111
 Db 540 TTGCAGGCTTCTTCTCCGATCATCAAGGCTGCCCCCTCTCTGTCATCATGATCAGCA 599
 QY 1112 CTTATGAGTTCGGCAAAAGCTTCTTCCAGAGCTGAACCCAGGACCGGCTTCTGGCGGCT 1171
 Db 600 CTTACGAGTTTGGCAAAAGCTTCTTCCAGAGCTGAACCCAGGACCGGCTTCTGGCGGCT 659
 QY 1172 GAAAGGCGCAAGCA 1185
 Db 660 TGAAGTGGGAGCA 673

RESULT 13
 ABN94635/c
 ID ABN94635 standard; DNA; 452 BP.
 XX AC ABN94635;
 XX DT 13-AUG-2002 (first entry)
 XX DE Gene #1133 used to diagnose liver cancer.
 XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
 XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX PN WO200229103-A2.
 XX PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US30589.
 XX PR 02-OCT-2000; 2000US-237054P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX DR WPI; 2002-426119/45.
 XX EX Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX

Db 457 AAGATTGCGCGATGAGGCGCTAGGACCTGTNNANNGGCTCCNAGCCACCTGGNG 516
 QY 490 ATGACTGCGCAGTACCGCCATCTACTTCACTGCCTATGACCAACTGAAGCCCTTCTG 549
 Db 517 ATGACGCGGCGAGTACTGCTATCTACTTCACTGCCTATGACCAACTGAGCCCTTCTG 576
 QY 550 TGTGCTCGACGCTGACCTTCACTTCACTGACGACCCATGCTGCGCGCTGCGCCG 609
 Db 577 TGTGG-CAGTCCTTGACCTCTGACCTTACGACCCATGCTGCGCGCTGCGCCG 635
 QY 610 CTGGGACCGCTGACTGTGATCAGCCCTGAGCTTATGGGACAAAGCTGAGGCTCAG 669
 Db 636 ATGGGACCGCTGACCT-TGGAGCCCTTGNAGCTGTGGGACCAAGCTGAGGCTCAG 694
 QY 670 CATGTGCTGACCGGAGTGGTGGCTGTGTTTCAACTGCACTGCTGAGGCTG 724
 Db 695 CATGTGCTGACCGTGGCTGCTTCTCTTT-AACTGCGNGAGCTCAGGCTG 748
 RESULT 15
 AAC76691/c
 ID AAC76691 standard; cDNA; 6712 BP.
 AC AAC76691;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 XX Human ORFX ORF2246 polynucleotide sequence SEQ ID NO:4491.
 DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antiprosoratic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunosulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antihypertensive;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; ant inflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX Homo sapiens.
 OS
 QX WO200058473-A2.
 FN
 PN
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Shimketa RA, Leach M;
 PI
 XX MPI; 2000-602362/57.
 DR P-PSDB; AAB42482.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 XX useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5; Page 3671-3675; 5507pp; English.
 PS
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;

CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunosulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 6712 BP; 2319 A; 1206 C; 1244 G; 1942 T; 1 other;
 Query Match 15.7%; Score 242.2; DB 21; Length 6712;
 Best Local Similarity 56.2%; Pred. No. 1.4e-49;
 Matches 476; Conservative 0; Mismatches 368; Indels 3; Gaps 1;
 QY 302 CCAAAATGGAAGTCCCTCTGTATTGCAATGCTGCTGGAGCTCTGTACCTGTGCCAA 361
 Db 3568 CCAAGGAAATGTTTGTATATAGTATGATCATGATCATCTATGTCTGTGAAG 3509
 QY 362 ATGTGCGCCCTGTGCGACCTGGTTTCAAGACCTTACCTGGCACCATTGATG 421
 Db 3508 AGGGAG---GCAACAACTATGTTATAGAGGCGAGAAATTTCCAGGGAACATGATG 3452
 QY 422 CTTTCGGAAGTCGAGGACGAGGCGCCACGAGCCCTCTGGAGCGGCTCCCGGCCA 481
 Db 3451 CATTTTAAATCATTCGAATGAGGCGCAATTAATCTCTATGGAGTGGCTTCTCTCTA 3392
 QY 482 CCCTGTGTGATGCTGCGACCTACCCCTCTACTTCTACTGCTTATGACCACTGAAG 541
 Db 3391 CCCTAGTGTGCGAGTCTCTCCACAGCTATTTATTTTACCTGTATGATGATTAAGT 3332
 QY 542 CTTTCTGTGTGTGAGCCCTGACCTCTGACCTCTACGACCCCATGCTGGCTGGCGCGC 601
 Db 3331 CTCTCTGAGATTAAGTTAGGAGAAATGAAACCTGCATACCAATTTCTCTGGAATTG 3272
 QY 602 TGGCCCGCTGGGCGACCTGTGATCAGCCCTGGAGCTTATGCGGCAAGAGCTGC 661
 Db 3271 TAGCCAGATTGTCGAGTAACTGTGATAAGTCCACTAGATTGATTAGAACAGATGC 3212
 QY 662 AGGCTCAGCATGTGTGACCGGAGCTGGGTGCTGCTTCTGAACTGCACTGGCTCAGG 721
 Db 3211 AGTCCAAGAAATTTCTTACGTGGAATGCTGATGATTTGTGAGAAATCTGGAAG 3152
 QY 722 GTGGCTGGGCTCACTGTGGCTGGGCGCCACCTGCTGGAGATGTGCTTCTTCTG 781
 Db 3151 ATGGTTGGATTTCCCTTTGGAGGGGCTGGGCTGCTACTGTTTAGAGATGCTACTTCT 3092
 QY 782 CAGCCCTGTACTGTTCAACTATGAGCTGAGAGCTGGCTCAATGGGTTTCAAGCCGA 841
 Db 3091 CAGCATGTACTGTATCACTATGAAATTTTAAAGAGTGGTTATGTGAGAAATCTGGTT 3032
 QY 842 AGGACAGACTTCTGTGGGCATGAGCTTTGTGGCTGTGGCATCTCAGGAGCGGTGGCTG 901
 Db 3031 TATATCAGCAACATTTTATGATCACTTTTACCTCAGGGGCAATTTGCTGTTCTTCTG 2972
 QY 902 CAGTGTGCTACTACCTTTGACGTGTAAAGACCAAGCCAGCTGCTGCGGAGCGA 961
 Db 2971 CTGTTCACACTTTTACCATTTGATGTAGTAAACACAAAGAGCAGACAACTTTGACAT 2912
 QY 962 TGGAGGCTGTGAGAGTGAACCCCTGCTGATGTGAGCTTCACCTGCTGCTGCGGAGA 1021
 Db 2911 ATGAAAGTCATAAAATTTCTATGCTTTTGTGATATGTCAACCTGGATTAATAAGAGACA 2852

Thu Jan 23 09:11:33 2003

us-09-888-358-2.rng

Page 15

Qy 1022 TCCGGGCGAGTCGGGACCAAGGACTCTTTCAGGCTTCCTTCGGATCATCAAGG 1081
| | | | |
Db 2851 TTGTTGCTAAATAATGATTTTCGGATTATTTTCAGGCTAATTCCTCGCTTAATTAAAA 2792
| | | | |
Qy 1082 CTGCCCCCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGA 1141
| | | | |
Db 2791 TTGCTCCTGCTGTGCCAATTATGATCAGTACATATGAATTTGGAAAGGCTTTTTCAGA 2732
| | | | |
Qy 1142 GGCTGAA 1148
| | |
Db 2731 AACAAAA 2725
| | |

Search completed: January 22, 2003, 19:04:05
Job time : 387.534 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:32:11 ; Search time 63.3511 Seconds
(without alignments)
7484.038 Million cell updates/sec

Title: US-09-888-358-2
Perfect score: 1546
Sequence: 1 ggctagtgctgctgacg.....gtctaagatgatgaattc 1546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

To number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1443.8	93.4	1643	2	US-08-933-750C-68
2	1443.8	93.4	1643	3	US-09-234-613-68
3	522	33.8	537	4	US-09-605-785-387
4	522	33.8	537	4	US-09-439-313-387
5	522	33.8	537	4	US-09-352-616A-387
6	61	3.9	997	3	US-09-188-930-23
7	61	3.9	1816	3	US-09-188-930-262
8	48.8	3.2	1311	4	US-09-068-140A-9
9	48.8	3.2	5150	4	US-09-068-140A-14
10	43.2	2.8	7218	1	US-08-232-463-14
11	42.2	2.7	1594	2	US-08-933-750C-61
12	42.2	2.7	1594	3	US-09-234-613-61
13	40.8	2.6	1259	3	US-08-961-871-11
14	39.8	2.6	1192	4	US-09-142-565-1
15	39.6	2.6	897	4	US-09-434-288-6
16	39.2	2.5	1177	3	US-08-961-871-9
17	38.4	2.5	34303	2	US-08-735-609-4
18	38.4	2.5	34303	2	US-08-735-609-4
19	38.4	2.5	34303	3	US-09-315-372-4
20	38.4	2.5	34303	3	US-09-244-752-4
21	38.4	2.5	34303	3	US-09-245-497-4
22	38.4	2.5	34303	4	US-09-562-919-4
23	38.4	2.5	34382	2	US-08-374-483-6
24	38.4	2.5	35408	4	US-08-973-334-3
25	38.4	2.5	35408	4	US-09-563-869A-3
26	38.4	2.5	35408	4	US-08-549-489-3
27	38.4	2.5	35935	2	US-08-735-609-1

28 38.4 2.5 35935 2 US-08-735-609-1 Sequence 1, Appl
29 38.4 2.5 35935 3 US-08-379-452-43 Sequence 43, Appl
30 38.4 2.5 35935 3 US-09-315-372-1 Sequence 1, Appl
31 38.4 2.5 35935 3 US-09-244-752-1 Sequence 1, Appl
32 38.4 2.5 35935 3 US-09-245-497-1 Sequence 1, Appl
33 38.4 2.5 35935 3 US-09-409-670-43 Sequence 43, Appl
34 38.4 2.5 35935 4 US-08-562-919-1 Sequence 1, Appl
35 37.4 2.4 2634 1 US-08-196-218-31 Sequence 31, Appl
36 37.4 2.4 2634 1 US-08-681-953-31 Sequence 31, Appl
37 36.8 2.4 8460 1 US-08-469-005A-9 Sequence 9, Appl
38 35.8 2.3 793 4 US-08-998-416-471 Sequence 471, Appl
39 35.6 2.3 1535 1 US-08-910-973-10 Sequence 10, Appl
40 35.6 2.3 1535 4 US-09-499-227-10 Sequence 10, Appl
41 35.6 2.3 1550 3 US-09-234-332-3 Sequence 3, Appl
42 35.6 2.3 1722 4 US-09-385-028-15 Sequence 15, Appl
43 35.6 2.3 11604 4 US-09-385-028-13 Sequence 13, Appl
44 35.6 2.3 15079 4 US-09-385-028-1 Sequence 1, Appl
45 35.4 2.3 606 2 US-08-432-871C-48 Sequence 48, Appl

ALIGNMENTS

RESULT 1
US-08-933-750C-68
; Sequence 68, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNCOAT01

CLONE: 724157
US-08-933-750C-68

Query Match 93.4%; Score 1443.8; DB 2; Length 1643;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 100 CAGGACTGAAGCTTCAAGATGCTGACAGGACCCCTCGGGCATCAGCCCCCTCAGCAA 159
DB 187 CAGGACTGAAGCTTCAAGATGCTGACAGGACCCCTCGGGCATCAGCCCCCTCAGCAA 246
QY 160 ATGFTGSCCTCAGGACCCCGGCTGTGTACTCTCTTTCATGACACCCCTGGAGTG 219
DB 247 ATGFTGSCCTCAGGACCCCGGCTGTGTACTCTCTTTCATGACACCCCTGGAGTG 306
QY 220 GTGAAGGTTTCGCTGAGTCTCAGGGGCTTCATGAGGACCCCTGATGCTTCTCTCC 279
DB 307 GTGAAGGTTTCGCTGAGTCTCAGGGGCTTCATGAGGACCCCTGATGCTTCTCTCC 366
QY 280 AGACTGTGAGGCTCTCTCTATACCAATGGAAGTGTCTCTGATTGCAATGGTCTCTG 339
DB 367 AGACTGTGAGGCTCTCTCTATACCAATGGAAGTGTCTCTGATTGCAATGGTCTCTG 426
QY 340 GAGCCTCTGACTGTGCGCAAAATGGTGTCCCGCTGTGCCACCTGTTTCAAGACCTTACC 399
DB 427 GAGCCTCTGACTGTGCGCAAAATGGTGTCCCGCTGTGCCACCTGTTTCAAGACCTTACC 486
QY 400 CGCTTCACTGGCAGCATGATGCTGTGAAGATGTTGAGCAGCAGGCGCAGGAGCC 459
DB 487 CGCTTCACTGGCAGCATGATGCTGTGAAGATGTTGAGCAGCAGGCGCAGGAGCC 546
QY 460 CTCTGAGGCGCTCTCCCGCCACCTGCTGATGATGCTGCCAGCTACCGCATCTACTTC 519
DB 547 CTCTGAGGCGCTCTCCCGCCACCTGCTGATGATGCTGCCAGCTACCGCATCTACTTC 606
QY 520 ACTGCTATGACAACTGAAAGCCTTCTCTGTGTGAGCGCTGACCTGTGACCTCTAC 579
DB 607 ACTGCTATGACAACTGAAAGCCTTCTCTGTGTGAGCGCTGACCTGTGACCTCTAC 666
QY 580 GCACCATGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 639
DB 667 GCACCATGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 726
QY 640 GAGCTTATCGGACAAAGCTCAGGCTCAGCATGCTGTATACCGGAGCTGGGTGCTGT 699
DB 727 GAGCTTATCGGACAAAGCTCAGGCTCAGCATGCTGTATACCGGAGCTGGGTGCTGT 786
QY 700 GTTCGAACTGCACTGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
DB 787 GTTCGAACTGCACTGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
QY 760 GCCTTCAGAGTGTGCTTCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 819
DB 847 GCCTTCAGAGTGTGCTTCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
QY 820 TGCTCAATGGGTTTCAAGCCAGGACCAAGCTTCTGTGGGCTGAGCTTGTGGGCTGT 879
DB 907 TGCTCAATGGGTTTCAAGCCAGGACCAAGCTTCTGTGGGCTGAGCTTGTGGGCTGT 966
QY 880 GGCATCTCAGGACGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
DB 967 GGCATCTCAGGACGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
QY 940 CGCAGGCTGCTTGGGAGGATGAGGCTGTGAGGAGGACCCCTTGCATGTGAGCTCC 999
DB 1027 CGCAGGCTGCTTGGGAGGATGAGGCTGTGAGGAGGACCCCTTGCATGTGAGCTCC 1086
QY 1000 ACTCGCTGCTGCTGCGGAGGATCCGGGCGGAGTCCGGGACCAAGGAGCTTCTTTCAGGC 1059
DB 1087 ACTCGCTGCTGCTGCGGAGGATCCGGGCGGAGTCCGGGACCAAGGAGCTTCTTTCAGGC 1146
QY 1060 TTCCTTCTCGATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119

DB 1147 TTCCTTCTCGATCATCAAGCTGCCCCCTCTCTGTCATCATCAGACCTATGAG 1206
QY 1120 TTCGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTTCTGGGCGGCTGAAAGGG 1179
DB 1207 TTCGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTTCTGGGCGGCTGAAAGGG 1266
QY 1180 CAAGGAGCAAGGACCCCTCTCTCCACGATGGGAGAGGSCAGGAGGAGCCAGCC 1239
DB 1267 CAAGGAGCAAGGACCCCTCTCTCCACGATGGGAGAGGSCAGGAGGAGCCAGCC 1326
QY 1240 AAGTGCCTTCTTCTCAGCTCAGGAGGAGGCTTGTTCCTTCCCTCCCGGCAAG 1299
DB 1327 AAGTGCCTTCTTCTCAGCTCAGGAGGAGGCTTGTTCCTTCCCTCCCGGCAAG 1386
QY 1300 CTCAGGAGGAGGCTGTCCTCTCTGGGCGGCGCAGCCTTCTCAGACAACTTCTTCC 1359
DB 1387 CTCAGGAGGAGGCTGTCCTCTCTGGGCGGCGCAGCCTTCTCAGACAACTTCTTCC 1446
QY 1360 GCTGCTCAGTCTGTTGGGATCATCACTTACCCACCCCGGCAAGTTCAAGACCAATCTTCC 1419
DB 1447 GCTGCTCAGTCTGTTGGGATCATCACTTACCCACCCCGGCAAGTTCAAGACCAATCTTCC 1506
QY 1420 AGCTGCCCCCTCTGCTTCTTCCCTGTGTTGCTGTAGCTGGGATGTTCTCAGGAAACCAAG 1479
DB 1507 AGCTGCCCCCTCTGCTTCTTCCCTGTGTTGCTGTAGCTGGGATGTTCTCAGGAAACCAAG 1566
QY 1480 AAGCCCTCAGCTGCTGTAGTCTCCTCTGACCTTGTAAATCTTAAAGTCTAAAGATGAT 1539
DB 1567 AAGCCCTCAGCTGCTGTAGTCTCCTCTGACCTTGTAAATCTTAAAGTCTAAAGATGAT 1626
QY 1540 GAATTC 1546
DB 1627 GAATTC 1633

RESULT 2
US-09-234-613-68
; Sequence 68, Application US/09234613
; Patent No. 613,973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 68:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1643 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: SYN0AT01
 CLONE: 724157
 US-09-234-613-68

Query Match 93.4%; Score 1443.8; DB 3; Length 1643;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 100 CAGGACTGAGCTTCAAGATGGCTGACAGGACCTTGGGGGATCAGCCCTCCAGCAA 159
 DB 187 CAGGACTGAGCTTCAAGATGGCTGACAGGACCTTGGGGGATCAGCCCTCCAGCAA 246
 QY 160 ATGCTGGCTCAGGACCGGGCTGTGTTTACCTCTCTTATGACACCTTGGAGCTG 219
 DB 247 ATGCTGGCTCAGGACCGGGCTGTGTTTACCTCTCTTATGACACCTTGGAGCTG 306
 QY 220 GTGAGGTTGGCTTGCATGTCTAGCGCCCTCCATGGCCAGCGACTGATGCCCTTCTCC 279
 DB 307 GTGAGGTTGGCTTGCATGTCTAGCGCCCTCCATGGCCAGCGACTGATGCCCTTCTCC 366
 QY 280 AGACTGTGAGGCTTCTCTATACCAATGGAAGTGCCTCTGTATTGCAATGTTGCTG 339
 DB 367 AGACTGTGAGGCTTCTCTATACCAATGGAAGTGCCTCTGTATTGCAATGTTGCTG 426
 QY 340 GAGCTCTGTACTGTGCGCCAAATGTTGCGCCCTGTGCGCCCTGTGCGCCCTGTGCGCC 399
 DB 427 GAGCTCTGTACTGTGCGCCAAATGTTGCGCCCTGTGCGCCCTGTGCGCCCTGTGCGCC 486
 QY 400 CCGTTCACTGGGACCATGATGCTTCTGTGGAAGATCGTGAGGACAGGGGACAGGACC 459
 DB 487 CCGTTCACTGGGACCATGATGCTTCTGTGGAAGATCGTGAGGACAGGGGACAGGACC 546
 QY 460 CTCTGAGGCGCTCCCGGACCTTGTGATGACTGTGCGAGTACCGGCACTTACTTTC 519
 DB 547 CTCTGAGGCGCTCCCGGACCTTGTGATGACTGTGCGAGTACCGGCACTTACTTTC 606
 QY 520 ACTGCTATGACCACTGAAGGCTTCTCTGTGTTGAGGCTGAGGACAGGGGACAGGACC 579
 DB 607 ACTGCTATGACCACTGAAGGCTTCTCTGTGTTGAGGCTGAGGACAGGGGACAGGACC 666
 QY 580 GCACCATGTTGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTG 639
 DB 667 GCACCATGTTGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTG 726
 QY 640 GAGCTTATGGGCAAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 699
 DB 727 GAGCTTATGGGCAAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 786
 QY 700 GTTCGAATCGAGTGGCTCAGGCTGGCTGAGGCTGGCTGAGGCTGGCTGAGGCTGGCTG 759
 DB 787 GTTCGAATCGAGTGGCTCAGGCTGGCTGAGGCTGGCTGAGGCTGGCTGAGGCTGGCTG 846
 QY 760 GCCCTTCGAGTGGCTTCTCAGCCTCTGAGGCTGGCTGAGGCTGGCTGAGGCTGGCTG 819
 DB 847 GCCCTTCGAGTGGCTTCTCAGCCTCTGAGGCTGGCTGAGGCTGGCTGAGGCTGGCTG 906
 QY 820 TGGCTCAATGGCTTCAAGCGGACGACGACTTCTGTGGGATGAGCTTGTGGCTGGT 879
 DB 907 TGGCTCAATGGCTTCAAGCGGACGACGACTTCTGTGGGATGAGCTTGTGGCTGGT 966
 QY 880 GGCATCTCAGGACGGTGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939

DB 967 GGCATCTCAGGACGGTGGCTGAGTGTGACTTACCTTTGACGTGTAAGACCCAA 1026
 QY 940 CGCCAGTCTCTCTGGGACCGATGAGGCTGTGAGAGTGAACCCCTGATGTGGACTCC 999
 DB 1027 CGCCAGTCTCTCTGGGACCGATGAGGCTGTGAGAGTGAACCCCTGATGTGGACTCC 1086
 QY 1000 ACCTGGCTGTCTGGGAGGATCGGCGGAGTGGGGACCAAGGACTCTTTGACGGC 1059
 DB 1087 ACCTGGCTGTCTGGGAGGATCGGCGGAGTGGGGACCAAGGACTCTTTGACGGC 1146
 QY 1060 TTCTTCTCGGATCATCAAGGCTGCCCCCTCTCTGTCCTCATGATCAGCACCTATGAG 1119
 DB 1147 TTCTTCTCGGATCATCAAGGCTGCCCCCTCTCTGTCCTCATGATCAGCACCTATGAG 1206
 QY 1120 TTGGGCAAGGCTTCTTCAGAGGCTGAACAGGACCGGCTTCTGGGCGGTGAAGGGG 1179
 DB 1207 TTGGGCAAGGCTTCTTCAGAGGCTGAACAGGACCGGCTTCTGGGCGGTGAAGGGG 1266
 QY 1180 CAAGGAGGCAAGGACCCCGTCTCTCCACGATGGGGAGGGGAGGAGAGCCAGCC 1239
 DB 1267 CAAGGAGGCAAGGACCCCGTCTCTCCACGATGGGGAGGGGAGGAGAGCCAGCC 1326
 QY 1240 AAGTGGCTTTTCTCAGACATGAGGAGGGGCTTGTTCCTTCCCTCCCGGCGACAAG 1299
 DB 1327 AAGTGGCTTTTCTCAGACATGAGGAGGGGCTTGTTCCTTCCCTCCCGGCGACAAG 1386
 QY 1300 CTCAGGAGGAGGCTTCTCTCTGGGGGCGCA~ACTTCTCAGACACAACCTTCTTCT 1359
 DB 1387 CTCAGGAGGAGGCTTCTCTCTGGGGGCGCA~ACTTCTCAGACACAACCTTCTTCT 1446
 QY 1360 GCTGCTCCAGTGTGGGATCATCAGCTTACCCACCCCGGAGGTTCAAGACCAATCTTC 1419
 DB 1447 GCTGCTCCAGTGTGGGATCATCAGCTTACCCACCCCGGAGGTTCAAGACCAATCTTC 1506
 QY 1420 AGTGCCCTTCTGTTTCTCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1479
 DB 1507 AGTGCCCTTCTGTTTCTCTGTTTCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1566
 QY 1480 AAGCCCTCAGCTGTGTAGTCTCCCTGACCCCTGACCCCTGTTAATTCCTTAAGTATAT 1539
 DB 1567 AAGCCCTCAGCTGTGTAGTCTCCCTGACCCCTGACCCCTGTTAATTCCTTAAGTATAT 1626
 QY 1540 GAACTTC 1546
 DB 1627 GAACTTC 1633

RESULT 3
 US-09-605-785-387
 ; Sequence 387, Application US/09605785
 ; Patent No. 6321716
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Devin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqi
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darriek
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Hepler, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C16
 ; CURRENT APPLICATION NUMBER: US/09/605,785

; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-605-785-387

Query Match 33.8%; Score 522; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1025 GGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 1144
Db 61 CCCCCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 120
QY 1145 TGAACGAGACCGGCTTCTGGGGCTTGAAGGGCAAGGAGGCAAGGACCCCGTCTCTC 1204
Db 121 TGAACGAGACCGGCTTCTGGGGCTTGAAGGGCAAGGAGGCAAGGACCCCGTCTCTC 180
QY 1205 CCACGATGGGAGAGGAGGAGAGCCAGCCAGCTTCTTCCTCAGACTGAGG 1264
Db 181 CCACGATGGGAGAGGAGGAGAGCCAGCCAGCTTCTTCCTCAGACTGAGG 240
QY 1265 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 1324
Db 241 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 300
QY 1325 GGGGCCAGCACTTCTTCAGACAACTTCTTCCTGCTGCTCAGTCTGGGATCATCA 1384
Db 301 GGGGCCAGCACTTCTTCAGACAACTTCTTCCTGCTGCTCAGTCTGGGATCATCA 360
QY 1385 CTTACCCACCCCAAGTCTCAAGCAAAATCTCCAGCTGCCCTTCGTTTCCCTGT 1444
Db 361 CTTACCCACCCCAAGTCTCAAGCAAAATCTTCCTGCTGCTCAGTCTGGGATCATCA 420
QY 1445 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCTTGTAAATCTTAACTTAAGTCTAAAGATGATGAACCTC 1546
Db 481 CTGACCTTGTAAATCTTAACTTAAGTCTAAAGATGATGAACCTC 522

RESULT 4
US-09-439-313-387
; Sequence 387, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-439-313-387

Query Match 33.8%; Score 522; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1025 GGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 1144
Db 61 CCCCCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 120
QY 1145 TGAACGAGACCGGCTTCTGGGGCTTGAAGGGCAAGGAGGCAAGGACCCCGTCTCTC 1204
Db 121 TGAACGAGACCGGCTTCTGGGGCTTGAAGGGCAAGGAGGCAAGGACCCCGTCTCTC 180
QY 1205 CCACGATGGGAGAGGAGGAGAGCCAGCCAGCTTCTTCCTCAGACTGAGG 1264
Db 181 CCACGATGGGAGAGGAGGAGAGCCAGCCAGCTTCTTCCTCAGACTGAGG 240
QY 1265 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 1324
Db 241 GAGGGGCTTGTTCCTTCCCTCCGGGCAAGCTCCAGGCGAGGCTGTCCCTCTGG 300
QY 1325 GGGGCCAGCACTTCTTCAGACAACTTCTTCCTGCTGCTCAGTCTGGGATCATCA 1384
Db 301 GGGGCCAGCACTTCTTCAGACAACTTCTTCCTGCTGCTCAGTCTGGGATCATCA 360
QY 1385 CTTACCCACCCCAAGTCTCAAGCAAAATCTTCAGCTGCCCTTCGTTTCCCTGT 1444
Db 361 CTTACCCACCCCAAGTCTCAAGCAAAATCTTCAGCTGCCCTTCGTTTCCCTGT 420
QY 1445 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTCTGTAGCTGGGATCTCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCTTGTAAATCTTAACTTAAGTCTAAAGATGATGAACCTC 1546
Db 481 CTGACCTTGTAAATCTTAACTTAAGTCTAAAGATGATGAACCTC 522

RESULT 5
US-09-152-616A-387
; Sequence 387, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-352-616A-387

Query Match 33.8%; Score 522; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.1e-126;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCAGTCGGGACCAAGGAGCTTTTGCAGGCTTCTTTCCTCGGATCATCAAGGCTG 1084
 Db 1 GGGCCAGTCGGGACCAAGGAGCTTTTGCAGGCTTCTTTCCTCGGATCATCAAGGCTG 60
 QY 1085 CCCCTCTCTGCTCCATCATGATCAGACCTATGATTCGGGAAAGCTTCTTCCAGAGGC 1144
 Db 61 CCCCTCTCTGCTCCATCATGATCAGACCTATGATTCGGGAAAGCTTCTTCCAGAGGC 120
 QY 1145 TGAACAGGACCGGCTTCTGGCGGTGAAAGGGCAAGGAGGCAAGGACCCCGTCTTC 1204
 Db 121 TGAACAGGACCGGCTTCTGGCGGTGAAAGGGCAAGGAGGCAAGGACCCCGTCTTC 180
 QY 1205 CCACGATGGGAGAGGCGCAGGAGAGACCCAGCCAAAGTGCCTTTTCTCAGCATGAGG 1264
 Db 181 CCACGATGGGAGAGGCGCAGGAGAGACCCAGCCAAAGTGCCTTTTCTCAGCATGAGG 240
 QY 1265 GAGGGGCTGTTTCCCTTCCCTCCCGGCGACAGCTCCAGGAGGAGGCTGCTCCTG 1324
 Db 241 GAGGGGCTGTTTCCCTTCCCTCCCGGCGACAGCTCCAGGAGGAGGCTGCTCCTG 300
 QY 325 GGGGCCAGCACTTCTCAGACACAATCTTCTGCTGCTCCAGTCCGTTGGGATCATCA 1384
 Db 301 GGGGCCAGCACTTCTCAGACACAATCTTCTGCTGCTCCAGTCCGTTGGGATCATCA 360
 QY 1385 CTTACCCACCCCGCAAGTTCAGACCAATCTTCCAGTCCGCTTCTGTTTCCCTG 1444
 Db 361 CTTACCCACCCCGCAAGTTCAGACCAATCTTCCAGTCCGCTTCTGTTTCCCTG 420
 QY 1445 GTTCTGCTGAGTGGGATGCTCCAGGACCAAGAGCCCTCAGCTGCTGATGCTCC 1504
 Db 421 GTTCTGCTGAGTGGGATGCTCCAGGACCAAGAGCCCTCAGCTGCTGATGCTCC 480
 QY 1505 CTGACCTTGTAACTTCTTAAGTCTTAAGATGATGAATTC 1546
 Db 481 CTGACCTTGTAACTTCTTAAGTCTTAAGATGATGAATTC 522

RESULT 6
 US-09-188-930-23
 ; Sequence 23, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James Greg
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; FILE OF INVENTION: and Methods For Their Use
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: Fast-Seq for Windows Version 3.0
 ; SEQ ID NO 23
 ; LENGTH: 997
 ; TYPE: DNA
 ; ORGANISM: mouse
 US-09-188-930-23

Query Match 3.9%; Score 61; DB 3; Length 997;
 Best Local Similarity 52.1%; Pred. No. 1.1e-06;
 Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
 QY 393 CCCTACCGCTTCACTGGCACCATGATGCTTCTGTAAGATCGTGAGGACGAGGCGAC 452
 Db 718 CCGCAGCAACAATGTCATGCTAGGTGATTCACACAGATGATTCGAGAAGGGGAGC 777
 QY 453 CAGGACCTCTGGAGCGGCTCCCGCCACCTCGTGATGACTGTGCCAGCTACCGCCAT 512
 Db 778 CAAGTCACTCTGGCGGGCAAGGATCAATGTCCTCAAAATTCCTGAGTGGCCAT 837

QY 513 CTACTTCACTGCTATGACCACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTCTGA 572
 Db 838 CAATTCATGCGCATATGAGCAGATGAAACGGCTTGTGGTAGTATCAGGAGACGCTGAG 897
 QY 573 CTCTACGACCAACCATGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 632
 Db 898 GATCCACGAAAGGCTTGTGGCAGGCTCTTGGCCGAGGCCATTCGCCAGAGTAGCATCTA 957
 QY 633 CCCCTGGAGCTTATGCGGAC 653
 Db 958 CCAATGGAGTTCTGAAGAC 978

RESULT 7
 US-09-188-930-262
 ; Sequence 262, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James Greg
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; FILE OF INVENTION: and Methods For Their Use
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: Fast-Seq for Windows Version 3.0
 ; SEQ ID NO 262
 ; LENGTH: 1816
 ; TYPE: DNA
 ; ORGANISM: Mouse
 US-09-188-930-262

Query Match 3.9%; Score 61; DB 3; Length 1816;
 Best Local Similarity 52.1%; Pred. No. 1.4e-06;
 Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
 QY 393 CCCTACCGCTTCACTGGCACCATGATGCTTCTGTAAGATCGTGAGGACGAGGCGAC 452
 Db 718 CCGCAGCAACAATGTCATGCTAGGTGATTCACACAGATGATTCGAGAAGGGGAGC 777
 QY 453 CAGGACCTCTGGAGCGGCTCCCGCCACCTCGTGATGACTGTGCCAGCTACCGCCAT 512
 Db 778 CAAGTCACTCTGGCGGGCAAGGATCAATGTCCTCAAAATTCCTGAGTGGCCAT 837
 QY 513 CTACTTCACTGCTATGACCACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTCTGA 572
 Db 838 CAATTCATGCGCATATGAGCAGATGAAACGGCTTGTGGTAGTATCAGGAGACGCTGAG 897
 QY 573 CTCTACGACCAACCATGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 632
 Db 898 GATCCACGAAAGGCTTGTGGCAGGCTCTTGGCCGAGGCCATTCGCCAGAGTAGCATCTA 957
 QY 633 CCCCTGGAGCTTATGCGGAC 653
 Db 958 CCAATGGAGTTCTGAAGAC 978

RESULT 8
 US-09-068-140A-9
 ; Sequence 9, Application US/09068140A
 ; Patent No. 6281409
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
 ; APPLICANT: and Rex Michael Brennan
 ; TITLE OF INVENTION: Blackcurrant Promoters and Genes
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road

CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068,140A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04807
FILING DATE: No. 6281409ember 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: C70237
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Ribes nigrum
STRAIN: Ben Alder
US-09-068-140A-9

Query Match 3.2%; Score 48.8; DB 4; Length 1311;
Best Local Similarity 45.7%; Pred. No. 0.0019;
Matches 253; Conservative 0; Mismatches 292; Indels 9; Gaps 2;

QY 398 CCGCTTCACTGCGACCATGATGCTCTGTGAAGATCGTGAGGACGAGGCGACACAGA 457
Db 206 CTCAATCCCGCGCTCCGACAGCCCTTGGGTGCATCTGAAAGTTGAAGGTCGCCGCG 265
QY 458 CCCTCTGGAGCGGCTCCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCCATCTACT 517
Db 266 GACTTTACCGTGGCATTTGGTGAATGGTCTCGGTGCAGGACAGCTCAGCAGTGTATT 325
QY 518 TCAGTCCCTATGACCACTGAAGCCTTCTGTGTGTGTCGAGCCCTGACTCTGAGCCTCT 577
Db 326 TCTCCGTTTACGAGATGTGTAAAGAGACTTTTCTCATGGTGATCCGAGC---AATTCCG 382
QY 578 ACCACCATGCTGTGCTGGCGGCTGCGCGCTGGGACCGTGCATGCTGATCAGCCCC 637
Db 383 GTGCGGACCGCTTTCGGGGGTGTTCGACCGGTGCGAGCGGTCAGCGGTGATTAACCGCA 442
QY 638 TGGAGCTTATGCGGCAAAAGCTGAGGCTCAGCATGTGTCTGTACCGGAGCTGGTGCT 697
Db 443 TGGATGTGTGAACAGAGGTTGCAAGTTCGAGGACGATCCGTACAGGGTGTGTTGATT 502
QY 698 GTGTTCGAAGTCACTGAGTGGCTGAGGCTGCTGTGTGTGTCGAGCCCTGAGTGGGCCCCA 757
Db 503 GCGTGAAGAGGCTGTGTGAAGAGAGGATTTGGCGCATTTTACGCATCTTATCGAACAA 562
QY 758 CTGCGCTTCGAGATGTCCTTCTCAGCCCTGTACTGTTCAACTATGAGTGGTGAAGA 817
Db 563 CTGTGCTCATGATGATGCTTTCAGGCGGCTTCTGCGCACATATGAAGCAACAGAA 622
QY 818 GTTGGCTCAATGGGTTTCAGGCGGAA-----GGACGAGACTTCTGTGGGCATGAGTTTG 871
Db 623 AAGGTTGTGTGAGGTGTGCGCGGAGACTGCGAACGATGAGAAATTGTTAGTGCATGCTA 682

QY 872 TGGCTGTGTCATCTCAGGACGGTGGCTGAGTGTGACTTACCTTTGACGTGTAA 931
Db 693 CTGCTGTGCTGCTGCTGAGCTTTGGCTGAGTAGTAACCACTCCACTAGATTGTCA 742
QY 932 AGACCCAAAGCCAG 945
Db 743 AAACCTCAGTTGCAG 756

RESULT 9
US-09-068-140A-14
Sequence 14, Application US/09068140A
Patent No. 6281409
GENERAL INFORMATION:
APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
APPLICANT: and Rex Michael Brennan
TITLE OF INVENTION: Blackcurrant Promoters and Genes
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068,140A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/04807
FILING DATE: No. 6281409ember 4, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: C70237
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 5150 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Ribes nigrum
STRAIN: Ben Alder
US-09-068-140A-14

Query Match 3.2%; Score 48.8; DB 4; Length 5150;
Best Local Similarity 45.7%; Pred. No. 0.0032;
Matches 253; Conservative 0; Mismatches 292; Indels 9; Gaps 2;
QY 398 CCGCTTCACTGCGACCATGATGCTCTGTGAAGATCGTGAGGACGAGGCGACACAGA 457
Db 3370 CTCAATCCCGCGCTCCGACAGCCCTTGGGTGCATCTGAAAGTTGAAGGTCGCCGCG 3429
QY 458 CCCTCTGGAGCGGCTCCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCCATCTACT 517
Db 3430 GACTTTACCGTGGCATTTGGTGCATGGTCTCGGTGCAGGACAGCTCAGCAGTGTATT 3489
QY 518 TCAGTCCCTATGACCACTGAAGGCTTCTCTGTGTGTGTCGAGCCCTGACCTCTGACCTCT 577

TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 1594 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SPLNOT02
CLONE: 207452
US-08-933-750C-61

Query Match 2.7%; Score 42.2; DB 2; Length 1594;
Best Local Similarity 59.7%; Pred. No. 0.11;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1031 AGTCGGGACCAAGGACTCTTTCAGGCTTCCTCCGATCATCAAGGCTGCCCT 1090
DB 1001 AGAAGGCGCCCTGGCTTCTTCAAGGGCTGTCCCGAGCTTCTGAAGGCTGCCCT 1060
QY 1091 CCTGTCCATCATCATGACACCTATGATCGGCAAAAGCTTCTCCAGAGCTGAAC 1149
DB 1061 CCACAGGCTCATGTTCTTCTGTATGAATTTCTGTATGTTCTCCACTGATGAAC 1119

RESULT 12

US-09-234-613-61
Sequence 61, Application US/09234613
Patent No. 6132973

GENERAL INFORMATION:

APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234,613
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 1594 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SPLNOT02
CLONE: 207452
US-09-234-613-61

Query Match 2.7%; Score 42.2; DB 3; Length 1594;
Best Local Similarity 59.7%; Pred. No. 0.11;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1031 AGTCGGGACCAAGGACTCTTTCAGGCTTCCTCCGATCATCAAGGCTGCCCT 1090
DB 1001 AGAAGGCGCCCTGGCTTCTTCAAGGGCTGTCCCGAGCTTCTGAAGGCTGCCCT 1060
QY 1091 CCTGTCCATCATCATGACACCTATGATCGGCAAAAGCTTCTCCAGAGCTGAAC 1149
DB 1061 CCACAGGCTCATGTTCTTCTGTATGAATTTCTGTATGTTCTCCACTGATGAAC 1119

RESULT 13

US-08-961-871-11
Sequence 11, Application US/08961871
Patent No. 6013858

GENERAL INFORMATION:

APPLICANT: Wallace, Douglas C.
APPLICANT: Graham, Brett H.
APPLICANT: MacGregor, Grant R.
TITLE OF INVENTION: Mouse Lacking Heart-Muscle Adenine
Nucleotide Translocator Protein and Methods
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Wimmer and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,871
FILING DATE: 31-OCT-1997
CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/030,017
FILING DATE: 01-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 78-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

LENGTH: 1259 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 190..1086

US-08-961-871-11

Query Match 2.6%; Score 40.8; DB 3; Length 1259;
Best Local Similarity 50.0%; Pred. No. 0.22;
Matches 102; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 403 TTACTGGCACCATGATGCTCTGTGAAGATCGTGAGGACGAGGACACGAGCCCTC 462
Db 340 TACAAAGGCATCATTTGATGTGTGAGAAATCCCAAGGACGAGGCTTTCTCTCTTC 399
QY 463 TGGAGGGGCTCCCGCCACCTGCTGATGACTGTGCGAGCTACGCCATCTACTTACT 522
Db 400 TGGAGGGGTAACTGGCCAAAGTATCGGTACTTCCCACTCAAGCCCTGAATTCGCC 459
QY 523 GCCTATGACCACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTCTGACCTCTACGCA 582
Db 460 TTCAAAGACAGTACACAGCATTTCTCTGGAGGCTGATCGCCATAGCAGTTCTGG 519
QY 583 CCATGCTGGTGGCGCCCTGGCC 606
Db 520 CGTACTTTGCTGGTAACTGGCC 543

RESULT 14
US-09-142-565-1
; cience 1, Application US/09142565A
; nt No. 6187560
; ERAL INFORMATION:
; APPLICANT: Lee James Beeley
; APPLICANT: Kelly Paine
; APPLICANT: Robert James
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30002
; CURRENT APPLICATION NUMBER: US/09/142,565A
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: 9704551.2
; EARLIER FILING DATE: 1997-03-05
; EARLIER APPLICATION NUMBER: 9705614.7
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: 97305305.1
; EARLIER FILING DATE: 1997-07-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1192
; TYPE: DNA
; ORGANISM: HOMO SAPIEN
US-09-142-565-1

Query Match 2.6%; Score 39.8; DB 4; Length 1192;
Best Local Similarity 46.8%; Pred. No. 0.4;
Matches 196; Conservative 0; Mismatches 217; Indels 6; Gaps 2;

QY 403 TTCACTGGCACCATGATGCTCTGTGAAGATCGTGAGGACGAGGACACGAGCCCTC 462
DL 673 TACAGGGGACTATGGAGCGCTCAGAACCATCGCCAGGGAGGAGGTCAGGGGCTG 732
QY 463 TGGAGGGGCTCCCGCCACCTGCTGATGACTGTGCCAGCTACCGCATCTACTTCACT 522
Db 733 TGGAAAGGAACTTTGCCCAACATCATGAGGAATGCTATCGTCAACTGTGCTGAGGTGGT 792
QY 523 GCCTATGACCACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTCTGACCTCTACGCA 582
Db 793 ACCTAGCATCTCTCAAGGAGAGCTGCTGGACTACCACTGCTCACTGACAACTTCCCC 852
QY 583 ---CCCATGGTGGCTGGCGGCTGGCCCGCTGGCCACCGTGACTGATCAGCCCTG 639
Db 853 TGCCACTTTGCTCTGCTTTGGAGCGGCTTCTGTGCAAGTGGTGGCTTCCCGGCTG 912
QY 640 GAGCTTATCGGGAAGAGC---TGCAGGCTCAGCATGTGTGTCGACCGGAGCTGGGTGCC 696
Db 913 GACGTGGTGAAGACCGGCTATATGAATCACTCACTCCAGGCCAGTACTTACGCCCTTCAC 972
QY 697 TGTGTTGCACTGAGTGGCTCAGGTGGCTGGCGCTCACTGTGGGTGGGTGGGGCCCC 756
Db 973 TGTATGATAAGATGGTGGCCCCAGGGGCCCCACAGCGCTTCTACAAGGATTTACACC 1032
QY 757 ACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCACTATGAGCTGTGAA 815

Db 1033 TCCTTTTGGTTTGGATCCTTGGAACTGCTGATGTTCGTACCTATGACGAGCTGAA 1091

RESULT 15

US-09-434-288-6
; Sequence 6, Application US/09434288
; Patent No. 6303767
; GENERAL INFORMATION:
; APPLICANT: Betlach C., Melanie
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
; FILE REFERENCE: 30062-20030.00
; CURRENT APPLICATION NUMBER: US/09/434,288
; CURRENT FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: 60/107,093
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Streptomyces narbonensis
US-09-434-288-6

Query Match 2.6%; Score 39.6; DB 4; Length 897;
Best Local Similarity 47.2%; Pred. No. 0.4;
Matches 120; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 391 GACCTTACCCGCTTTCACCTGGCAACCATGGATGCTTCTGTAAGATCGTGAGGACGAGGGC 450
Db 295 GACCCCAACGACCCCAACCGGCTCACCGGCTGTGCGGAGCTCGCGACCTCGGGCG 354
QY 451 ACAGAGACCTCTGGAGGGGCTCCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCC 510
Db 355 ACGGCCACCGCTGCTGCTCTGCGACCTCACGACCGGGAGGCGCGCCGCTGCTCGCC 414
QY 511 ATCTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGGTGGAGCCCTGACTCT 570
Db 415 GGGGTCTCCGACGACGCCCTCAGGGCGCTCTCCACCTGCGGCCCAACCGTCGACTCC 474
QY 571 GACCTTACGCAACCATGGTGGCTGGCGGCTGCGCCGCTGGGACCGCTGACTGTGATC 630
Db 475 GAGCGCTCGCGCCACCGACCGCGAGCACTCGCCCGCTGTAACCGGAAGGCCACC 534
QY 631 AGCCCTCTGAGCT 644
Db 535 GCGGCGCTGCACT 548

Search completed: January 22, 2003, 19:06:08
Job time : 90.3511 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 17:33:27 ; Search time 77.3 Seconds

(without alignments)
8985.296 Million cell updates/sec

Title: US-09-888-358-2

Perfect score: 1546

Sequence: 1 ggctagggtcgctgcgagcg.....gtctaaagatgatgaattc 1546

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Tc number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1546	100.0	1546	10	US-09-888-358-2
2	1443.8	93.4	1643	10	US-09-840-787-68
3	1054.4	68.2	1114	10	US-09-888-358-1
4	522	33.8	537	9	US-10-012-896-387
5	522	33.8	537	9	US-09-895-793-387
6	522	33.8	537	9	US-09-895-814-387
7	522	33.8	537	10	US-09-759-143-387
8	522	33.8	537	10	US-09-780-669-387
9	522	33.8	537	10	US-09-822-827-387
10	445.4	28.8	452	10	US-09-880-107-1133
11	218.4	14.1	240	10	US-09-860-107-1179
12	185.8	12.0	453	10	US-09-867-701-4486
13	172	11.1	275	10	US-09-908-711-46
14	170.2	11.0	512	10	US-09-917-800A-329
15	161	10.4	246	10	US-09-960-352-2543
16	74.8	4.8	452	10	US-09-864-761-11674
17	74.8	4.8	496	10	US-09-864-761-11708
18	60	3.9	3334	9	US-09-982-598-288
19	60	3.9	3334	9	US-09-989-293A-288

20	60	3.9	3334	9	US-10-063-547-57	Sequence 57, Appl
21	60	3.9	3334	9	US-09-989-735-288	Sequence 288, Appl
22	60	3.9	3334	9	US-09-990-444-288	Sequence 288, Appl
23	60	3.9	3334	9	US-09-989-730-288	Sequence 288, Appl
24	60	3.9	3334	9	US-09-990-436-288	Sequence 288, Appl
25	60	3.9	3334	9	US-09-991-181-288	Sequence 288, Appl
26	60	3.9	3334	9	US-09-993-687-288	Sequence 288, Appl
27	60	3.9	3334	9	US-09-989-734-288	Sequence 288, Appl
28	60	3.9	3334	9	US-09-997-653-288	Sequence 288, Appl
29	60	3.9	3334	9	US-10-063-616-57	Sequence 57, Appl
30	60	3.9	3334	10	US-09-989-722-288	Sequence 288, Appl
31	60	3.9	3334	10	US-09-989-723-288	Sequence 288, Appl
32	60	3.9	3334	10	US-09-989-729-288	Sequence 288, Appl
33	60	3.9	3334	10	US-09-989-727-288	Sequence 288, Appl
34	60	3.9	3334	10	US-09-989-731-288	Sequence 288, Appl
35	60	3.9	3334	10	US-09-989-732-288	Sequence 288, Appl
36	60	3.9	3334	10	US-09-991-073-288	Sequence 288, Appl
37	60	3.9	3334	10	US-09-990-442-288	Sequence 288, Appl
38	60	3.9	3334	10	US-09-991-163-288	Sequence 288, Appl
39	60	3.9	3334	10	US-09-993-604-288	Sequence 288, Appl
40	60	3.9	3334	10	US-09-990-456-288	Sequence 288, Appl
41	60	3.9	3334	10	US-09-989-721-288	Sequence 288, Appl
42	60	3.9	3334	12	US-10-006-867-57	Sequence 57, Appl
43	57.6	3.7	144	10	US-09-864-761-28253	Sequence 28253, A
44	55.4	3.6	129	10	US-09-864-761-28284	Sequence 28284, A
45	50.6	3.3	3555	12	US-10-044-090-527	Sequence 527, Appl

ALIGNMENTS

RESULT 1
US-09-888-358-2
; Sequence 2, Application US/09888358
; Patent No. US20020119137A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID
; APPLICANT: ADAMS, SEAN H.
; APPLICANT: YU, XING XIAN
; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 10716/66
; CURRENT APPLICATION NUMBER: US/09/888,358
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,307
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-358-2

Query Match	100.0%	Score 1546;	DB 10;	Length 1546;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1546;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGCTAGGTGCGCTCGAGCGCGGAGCCACGAGGCGGACGACGTAATGGCCCGCCT	60	
Db	1	GGCTAGGTGCGCTCGAGCGCGGAGCCACGAGGCGGACGACGTAATGGCCCGCCT	60	
QY	61	GGCCCTGGGCGCGCGCGGCGGACGAGCCAGGCGGAGCTAGAGCAGGACTGAGCTTCAGATG	120	
Db	61	GGCCCTGGGCGCGCGCGGCGGACGAGCCAGGCGGAGCTAGAGCAGGACTGAGCTTCAGATG	120	
QY	121	GCTGACGAGGACCTTGGGGGATCAGCCCGCTCCAGCAATATGGTGGCTCAGGACCGGG	180	
Db	121	GCTGACGAGGACCTTGGGGGATCAGCCCGCTCCAGCAATATGGTGGCTCAGGACCGGG	180	
QY	181	GCTGTGGTTACTCTCTCTTCATGACACCCCTGAGAGGTTCCCTCGAGTCT	240	
Db	181	GCTGTGGTTACTCTCTCTTCATGACACCCCTGAGAGGTTCCCTCGAGTCT	240	

QY 241 CAGCGCCCTCCATGCGCAGCGAGCTGATGCTTCTCCAGAGCTGCGAGCTCTCCTAT 300
DB 241 CAGCGCCCTCCATGCGCAGCGAGCTGATGCTTCTCCAGAGCTGCGAGCTCTCCTAT 300
QY 301 ACCAAATGGAAGTGCCTCTCTGATTTGCAATAGTGTCTCTGAGCCTCTGTAAGTGCCTCA 360
DB 301 ACCAAATGGAAGTGCCTCTCTGATTTGCAATAGTGTCTCTGAGCCTCTGTAAGTGCCTCA 360
QY 361 AATGTTGCCCGCTGTCGACCTGTTTCAAGACCTTACCGCTTCACTGGCACCATGGAT 420
DB 361 AATGTTGCCCGCTGTCGACCTGTTTCAAGACCTTACCGCTTCACTGGCACCATGGAT 420
QY 421 GCCTTCTGTAAGATGCTGAGGACAGGACGAGGACCTCTGAGGCGGCTCCCGCC 480
DB 421 GCCTTCTGTAAGATGCTGAGGACAGGACGAGGACCTCTGAGGCGGCTCCCGCC 480
QY 481 ACCCTGATGATGCTGTCAGCTACCGCCATCTACTTCACTGCTTATGACCAACTGAAG 540
DB 481 ACCCTGATGATGCTGTCAGCTACCGCCATCTACTTCACTGCTTATGACCAACTGAAG 540
QY 541 GCCTTCTGTTGTCGAGCCTGACCTCTGACCTCTAGCACCCTAGTGTGCTGCGCGG 600
DB 541 GCCTTCTGTTGTCGAGCCTGACCTCTGACCTCTAGCACCCTAGTGTGCTGCGCGG 600
QY 601 CTGGCCCGCTGGCACCGTGAAGTGTGATCAGCCCTCTGAGCTTATGCGGACCAAGCTG 660
DB 601 CTGGCCCGCTGGCACCGTGAAGTGTGATCAGCCCTCTGAGCTTATGCGGACCAAGCTG 660
QY 661 CAGGCTCAGCATGTTGCTACCGGAGCTGGTGGCTGCTGCTGCAAGTGGCTCAG 720
DB 661 CAGGCTCAGCATGTTGCTACCGGAGCTGGTGGCTGCTGCTGCAAGTGGCTCAG 720
QY 721 GGTGGCTGGCTCACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 780
DB 721 GGTGGCTGGCTCACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 780
QY 781 TCAGCCCTGCTACTGCTTCAATATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGGCCG 840
DB 781 TCAGCCCTGCTACTGCTTCAATATGAGCTGGTGAAGAGCTGGCTCAATGGTTCAGGCCG 840
QY 841 AAGGACAGACTTGTGGGCAAGCTTTGTGGCTGGTGGCTTCTCAGGAGCGTGGCT 900
DB 841 AAGGACAGACTTGTGGGCAAGCTTTGTGGCTGGTGGCTTCTCAGGAGCGTGGCT 900
QY 901 GCAGTCTGACTTACCTTTGAGCTGGTGAAGACCCAGGAGCTGGCTTCTGGAGCG 960
DB 901 GCAGTCTGACTTACCTTTGAGCTGGTGAAGACCCAGGAGCTGGCTTCTGGAGCG 960
QY 961 ATGAGGCTGTGAGAGTGAACCCCTGCTGATGAGCTGCACTGCTGCTGCTGCTGGAGG 1020
DB 961 ATGAGGCTGTGAGAGTGAACCCCTGCTGATGAGCTGCACTGCTGCTGCTGCTGGAGG 1020
QY 1021 ATCCGGCGGAGTGGGCAAGAGGACTTTTTCAGGCTTCTTCTTCGATCATCAAG 1080
DB 1021 ATCCGGCGGAGTGGGCAAGAGGACTTTTTCAGGCTTCTTCTTCGATCATCAAG 1080
QY 1081 GCTGCCCCCTCTGTCATCATGATCAGCACTTATGATTTGGCAAAAGCTTCTTCAG 1140
DB 1081 GCTGCCCCCTCTGTCATCATGATCAGCACTTATGATTTGGCAAAAGCTTCTTCAG 1140
QY 1141 AGGCTGAACAGGACCGCTTCTGGCGGCTGAAGGGGCAAGGAGCAAGGACCCGCTC 1200
DB 1141 AGGCTGAACAGGACCGCTTCTGGCGGCTGAAGGGGCAAGGAGCAAGGACCCGCTC 1200
QY 1201 TCTCCACGATGGGAGAGGAGGAGGAGACCCAGGCAAGTGCCTTTTCTTCAGCACT 1260
DB 1201 TCTCCACGATGGGAGAGGAGGAGGAGACCCAGGCAAGTGCCTTTTCTTCAGCACT 1260
QY 1261 GAGGAGGGGGCTGTTTCTCCCTTCTCCCGGCAAGCTTCCAGGCGAGGCTGTCCCT 1320
DB 1261 GAGGAGGGGGCTGTTTCTCCCTTCTCCCGGCAAGCTTCCAGGCGAGGCTGTCCCT 1320
QY 1321 CTGGCGGCGCCAGCATTCCTCTCAGACACAATCTTCTTCTGCTGCTCCAGCTGCTGGGATC 1380

DB 1321 CTGGCGGCGCCAGCATTCCTCTCAGACACAATCTTCTTCTGCTGCTCCAGTCTGGGATC 1380
QY 1381 ATCACTTTACCCACCCCAAGTTCAAGACCAAAATCTTCCAGTGGCCCTTCTGTTTCC 1440
DB 1381 ATCACTTTACCCACCCCAAGTTCAAGACCAAAATCTTCCAGTGGCCCTTCTGTTTCC 1440
QY 1441 CTGCTTTGCTGCTAGTGGGATGCTTCCAGGAACCAAGACCCCTCAGCCTGCTGTTAGT 1500
DB 1441 CTGCTTTGCTGCTAGTGGGATGCTTCCAGGAACCAAGACCCCTCAGCCTGCTGTTAGT 1500
QY 1501 CTCCTGACCCCTGTTAATCTTAACTTAACTTAAAGATGATGAATTC 1546
DB 1501 CTCCTGACCCCTGTTAATCTTAACTTAACTTAAAGATGATGAATTC 1546

RESULT 2
US-09-840-787-68
; Sequence 68, Application US/09840787
; Patent No. US20020058264A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Bandman, Olga
; Shah, Purvi
; Au-Young, Janice
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/840,787
; FILING DATE: 23-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/518,865
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYN00AT01
; CLONE: 724157
; SEQUENCE DESCRIPTION: SEQ ID NO: 68 :
US-09-840-787-68

Query Match 93.4%; Score 1443.8; DB 10; Length 1643;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1445; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 100 CAGGACTGAAGCTTCAAGATGGCTGACACAGGACCTTGGGGCATACAGCCCTCCAGAA 159
Db 187 CAGGACTGAAGCTTCAAGATGGCTGACACAGGACCTTGGGGCATACAGCCCTCCAGAA 246
Qy 160 ATGGTGGCTCAGGACCGGGGTGGTTAACTCTCTTTTATGACACCCCTGGACGFG 219
Db 247 ATGGTGGCTCAGGACCGGGGTGGTTAACTCTCTTTTATGACACCCCTGGACGFG 306
Qy 220 GTGAGGTTCCGCTGCGATCTCAGCGGCTTCCATGGCGAGGCTGATGCCCTTCTCC 279
Db 307 GTGAGGTTCCGCTGCGATCTCAGCGGCTTCCATGGCGAGGCTGATGCCCTTCTCC 366
Qy 280 AGACTGTGAGCCTCTCTATACAAATGGAAGTGCCTCTGTATTCGAATGTGTCCTG 339
Db 367 AGACTGTGAGCCTCTCTATACAAATGGAAGTGCCTCTGTATTCGAATGTGTCCTG 426
Qy 340 GAGCCTGTGATCTGTGCGCAATGTGCGCGCTGTGCACTGGTTTCAAGACCCCTACC 399
Db 427 GAGCCTGTGATCTGTGCGCAATGTGCGCGCTGTGCACTGGTTTCAAGACCCCTACC 486
Qy 400 CCGTCACTGGCACCATGATGCTTCTGTGAAGATCGTGAGGACAGGAGGACAGGAC 459
Db 487 CCGTCACTGGCACCATGATGCTTCTGTGAAGATCGTGAGGACAGGAGGACAGGAC 546
Qy 460 CTCTGAGCGGCTCCCGGCCACCCCTGTGATGATGTGCCAGTACCGCCATCTACTTC 519
Db 547 CTCTGAGCGGCTCCCGGCCACCCCTGTGATGATGTGCCAGTACCGCCATCTACTTC 606
Qy 520 ACTGCCATGACCAACTGAAGGCTTCTGTGTGTGAGCCCTGAGCCCTGACCTGTAC 579
Db 607 ACTGCCATGACCAACTGAAGGCTTCTGTGTGTGAGCCCTGAGCCCTGACCTGTAC 666
Qy 580 GCACCATGTGTGGTGGCGCTGGCCGCTGGGACCGTGATGATGATGATGATGATGAT 639
Db 667 GCACCATGTGTGGTGGCGCTGGCCGCTGGGACCGTGATGATGATGATGATGATGAT 726
Qy 640 GAGCTTATCGGCAAAAGCTGCAGGCTCAGAGTGTGTGTACCGGGAGCTGGTGCTGT 699
Db 727 GAGCTTATCGGCAAAAGCTGCAGGCTCAGAGTGTGTGTACCGGGAGCTGGTGCTGT 786
Qy 700 GTTCGAATGAGTGGCTCAGGCTGGCTGGCTCACTGTGGCTGGCTGGGGCCCCACT 759
Db 787 GTTCGAATGAGTGGCTCAGGCTGGCTGGCTCACTGTGGCTGGCTGGGGCCCCACT 846
Qy 760 GCGCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAACTATGATGATGATGATGAT 819
Db 847 GCGCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAACTATGATGATGATGATGAT 906
Qy 820 TGGCTCAATGGGTTTCAAGGACAGGACCACTTCTGTGGCATGAGCTTTTGTGGCTGT 879
Db 907 TGGCTCAATGGGTTTCAAGGACAGGACCACTTCTGTGGCATGAGCTTTTGTGGCTGT 966
Qy 880 GGCATCTCAGGACGAGTGGCTGAGTGTGATCTACCTTTGAGTGTGAAGACCCAA 939
Db 967 GGCATCTCAGGACGAGTGGCTGAGTGTGATCTACCTTTGAGTGTGAAGACCCAA 1026
Qy 940 CCGCAGGTCGCTTCTGGGAGCGATGGAGGCTGTGAGTGAACCCCTGCATGTGAGCTCC 999
Db 1027 CCGCAGGTCGCTTCTGGGAGCGATGGAGGCTGTGAGTGAACCCCTGCATGTGAGCTCC 1086
Qy 1000 ACTGGCTGTGCTGCGAGGATCCGGGCGGAGTCCGGGACCAAGGAGCTCTTTTGCAGGC 1059
Db 1087 ACTGGCTGTGCTGCGAGGATCCGGGCGGAGTCCGGGACCAAGGAGCTCTTTTGCAGGC 1146
Qy 1060 TTCCCTTCTCGGATCATCAAGGCTGCCCTCTCTGTGCATCATGATCAGACCTATGAG 1119
Db 1147 TTCCCTTCTCGGATCATCAAGGCTGCCCTCTCTGTGCATCATGATCAGACCTATGAG 1206
Qy 1120 TTCCGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTTCTGGCGGCTGAAGGGG 1179
Db 1207 TTCCGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTTCTGGCGGCTGAAGGGG 1266

Qy 1180 CAAGAGGCAAGGACCCCGTCTCTCCCAAGATGGGAGAGGAGGAGGAGGAGGAGGAGG 1239
Db 1267 CAAGAGGCAAGGACCCCGTCTCTCCCAAGATGGGAGAGGAGGAGGAGGAGGAGGAGG 1326
Qy 1240 AAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1299
Db 1327 AAGTGCCTTTTCTCAGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1386
Qy 1300 CTCAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1359
Db 1387 CTCAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1446
Qy 1360 GCTGCTCCAGTCTGGGAGTATCATCTTACCCACCCCAAGATTCAGAGCAAAATCTTCC 1419
Db 1447 GCTGCTCCAGTCTGGGAGTATCATCTTACCCACCCCAAGATTCAGAGCAAAATCTTCC 1506
Qy 1420 AGTGCCTTCTGCTGTTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1479
Db 1507 AGTGCCTTCTGCTGTTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1566
Qy 1480 AAGCCCTCAGGCTGT 1539
Db 1567 AAGCCCTCAGGCTGT 1626
Qy 1540 GAACTTC 1546
Db 1627 GAACTTC 1633
RESULT 3
US-09-888-358-1
; Sequence 1, Application US/09888358
; Patent No. US20020119137A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID
; APPLICANT: ADAMS, SEAN H.
; APPLICANT: YU, XING XIAN
; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 10716/66
; CURRENT APPLICATION NUMBER: US/09/888,358
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,307
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-358-1
Query Match 58.2%; Score 1054.4; DB 10; Length 1114;
Best Local Similarity 97.8%; Pred. No. 3.9e-274;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
Qy 105 CTGAAGCTTCAAGATGGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 164
Db 1 CTGAAGCTTCAAGATGGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
Qy 165 GGCCTCAGGACCGGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGACGTTGGTGA 224
Db 61 GGCCTCAGGACCGGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGACGTTGGTGA 120
Qy 225 GGTTCGCTGCAAGTCTCAGGCGGCTTCATGCGGAGGAGGAGGAGGAGGAGGAGGAGG 284
Db 121 GGTTCGCTGCAAGTCTCAGGCGGCTTCATGCGGAGGAGGAGGAGGAGGAGGAGGAGG 180
Qy 285 GTGGAGCTCTCTTATACCAAT-----GGAGGTGCTCTCT 320
Db 181 GTGGAGCTCTCTTATACCAATTTGCCCTCTCTCTCCATCCACAGGAGGAGGAGGAGGAGG 240
Qy 321 GTATTGCAATGGTCTCTGGAGGCTCTGTACCTGTGCCAAATGGTCCCGCTGTGCCAC 380

Db	241	GTATTGCAATGGTGCTCTGGAGGCTCTGTACTGTGGCCCAATATGGTGGCCGCTGTGCCAC	300
QY	381	CTGGTTTCAAGACCCCTACCGCTTCACTGGCAACATGGATGCTTCGTGAAGATCGTGAG	440
Db	301	CTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATGGATGCTTCGTGAAGATCGTGAG	360
QY	441	GCACGAGGGACACAGGACCTCTGGAGCGGCTCCCGCCACCTGGTGATGACTGTGCC	500
Db	361	GCACGAGGGACACAGGACCTCTGGAGCGGCTCCCGCCACCTGGTGATGACTGTGCC	420
QY	501	AGTACCGCCATCTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGTCGAGC	560
Db	421	AGTACCGCCATCTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGTCGAGC	480
QY	561	CCTGACCTCTGAACTCTACGCCACCATGGTGGCTGGCGGCTGGCCGCTGGGCAACCGT	620
Db	481	CCTGACCTCTGAACTCTAGCACCCATGGTGGCTGGCGGCTGGCCGCTGGGCAACCGT	540
QY	621	GACTGTGATCAGCCCTCGAGCTTATCGGACAAAGCTTCGAGGCTCAGCATGTGCTGA	680
Db	541	GACTGTGATCAGCCCTCGAGCTTATCGGACAAAGCTTCGAGGCTCAGCATGTGCTGA	600
QY	681	CCGGAGCTGGGTGCTGTGTTCGAACTGCAGTGGCTCAGGGTGGCTGGCGCTCACTGTG	740
Db	601	CCGGAGCTGGGTGCTGTGTTCGAACTGCAGTGGCTCAGGGTGGCTGGCGCTCACTGTG	660
QY	741	GCTGGGCTGGGGCCCCACTGCCCCCTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTCAA	800
Db	661	GCTGGGCTGGGGCCCCACTGCCCCCTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTCAA	720
QY	801	CTATCAGCTGCTGAAGAGCTGGCTCAATGGGTTTCAGGCCCGAGGACCAAGACTTCTGTGGG	860
Db	721	CTATCAGCTGCTGAAGAGCTGGCTCAATGGGCTCAGGCCCGAGGACCAAGACTTCTGTGGG	780
QY	861	CATGACTTTTGGCTGGTGGGATCTCAGGGAACGGTGGCTGCAGTGCCTGACTCTACCCCT	920
Db	781	CATGACTTTTGGCTGGTGGGATCTCAGGGAACGGTGGCTGCAGTGCCTGACTCTACCCCT	840
QY	921	TCAGCTGCTAAAGACCCACGCCAGTGCCTCGGAGCGATCGGAGGCTGTGAGAGTGAA	980
Db	841	TCAGCTGCTAAAGACCCACGCCAGTGCCTCGGAGCGATCGGAGGCTGTGAGAGTGAA	900
QY	981	CCCCCTGCAATGTGACTTCCACCTGGCTCTGCTCGGAGGATCCGGGGCCGAGTGGGGCAC	1040
Db	901	CCCCCTGCAATGTGACTTCCACCTGGCTCTGCTCGGAGGATCCGGGGCCGAGTGGGGCAC	960
QY	1041	GAAGGGACTTTTGCAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCCCTCTGTGTGCAT	1100
Db	961	GAAGGGACTTTTGCAGGCTTCTTCTCTCGGATCATCAAGGCTGCCCCCTCTGTGTGCAT	1020
QY	1101	CATGATCAGCACTATGAGTTGGCAAAAGCTTCTTCAGAGGCTGAACCAAGGACCGGCT	1160
Db	1021	CATGATCAGCACTATGAGTTGGCAAAAGCTTCTTCAGAGGCTGAACCAAGGACCGGCT	1080
QY	1161	CTGGCGGCTGAAAGGGGCAAGGGCAAGGAC	1194
Db	1081	CTGGCGGCTGAAAGGGGCAAGGGCAAGGAC	1114

RESULT 4

US-10-012-896-387

US-10-012-898-387
: Sequence 387. Application US/10012896

Sequence 387, Application US/10
Publication No. US20020183251A1

; PUBLICATION NO. 95200
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: Xu Jiaochun

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C

APPLICANT: DILLON, Davin C.
APPLICANT: MITCHEM, Jennifer L.

APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan

APPLICANT: Jiang, Yugu

APPLICANT: Kałos, Michael

; APPLICANT: Retter, Marc W

APPLICANT: Stolk, John A.

```

/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Bassols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Wantanabe, Yoshihiro
/ APPLICANT: Meagher, Madeleine Joy
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
/ FILE REFERENCE: 210121.427C27
/ CURRENT APPLICATION NUMBER: US/10/012,896
/ CURRENT FILING DATE: 2001-12-10
/ NUMBER OF SEQ ID NOS: 1011
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 387
/ LENGTH: 537
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-012-896-387

```

5 JUL 1954

RESULTS
US-09-895-793-387

US=09-895-193-38 / . sequence 387 Application IIS/09895793

; sequence 387, Application US/09
: Publication No. US20020192763A1

GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-387

Query Match 33.8%; Score 522; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTCTCGGATCATCAAGGCTG 60

QY 1085 CCCCTCTCTGTCCTCATCATGATCAGACCTATGAGTTGGGAAAAGCTTCTTCAGAGGC 1144
Db 61 CCCCTCTCTGTCCTCATCATGATCAGACCTATGAGTTGGGAAAAGCTTCTTCAGAGGC 120

C 1145 TGAACAGGACCGGCTTCTGGGGGCTGAAAGGGGCAAGGAGGCAAGGACCCGCTCTCTC 1204
Db 121 TGAACAGGACCGGCTTCTGGGGGCTGAAAGGGGCAAGGAGGCAAGGACCCGCTCTCTC 180

QY 1205 CCAGGATGGGAGAGGAGGAGAGAGCCAGCCAGTGCCTTCTCAGCACTGAGG 1264
Db 181 CCAGGATGGGAGAGGAGGAGAGAGCCAGCCAGTGCCTTCTCAGCACTGAGG 240

QY 1265 GAGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGAGGCGTCCCTCTCG 1324
Db 241 GAGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGAGGCGTCCCTCTCG 300

QY 1325 GCGGCCAGCACTTCTCAGACAACTTCTCTGCTGCTCCAGTCTGGGATCATCA 1384
Db 301 GCGGCCAGCACTTCTCAGACAACTTCTCTGCTGCTCCAGTCTGGGATCATCA 360

QY 1385 CTTACCCACCCGAGTTCAAGACCAATCTTCAGCTGCGCCCTCTGTTTCCCTGT 1444
Db 361 CTTACCCACCCGAGTTCAAGACCAATCTTCAGCTGCGCCCTCTGTTTCCCTGT 420

QY 1445 GTTGTGCTAGTGGGATGCTCCAGAACCAAGAGCCCTCAGCTGGTGTAGTCTCC 1504
Db 421 GTTGTGCTAGTGGGATGCTCCAGAACCAAGAGCCCTCAGCTGGTGTAGTCTCC 480

QY 1505 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 1546

Db 481 CTGACCCCTTGTAAATTCCTTAAGTCTAAAGATGATGAATTC 522

RESULT 6
US-09-895-814-387
; Sequence 387, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-387

Query Match 33.8%; Score 522; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTCTCGGATCATCAAGGCTG 60

QY 1085 CCCCTCTCTGTCCTCATCATGATCAGACCTATGAGTTGGGAAAAGCTTCTTCAGAGGC 1144
Db 61 CCCCTCTCTGTCCTCATCATGATCAGACCTATGAGTTGGGAAAAGCTTCTTCAGAGGC 120

QY 1145 TGAACAGGACCGGCTTCTGGGGGCTGAAAGGGGCAAGGAGGCAAGGACCCGCTCTC 1204
Db 121 TGAACAGGACCGGCTTCTGGGGGCTGAAAGGGGCAAGGAGGCAAGGACCCGCTCTC 180

QY 1205 CCAGGATGGGAGAGGAGGAGAGAGCCAGCCAGTGCCTTCTCAGCACTGAGG 1264
Db 181 CCAGGATGGGAGAGGAGGAGAGAGCCAGCCAGTGCCTTCTCAGCACTGAGG 240

QY 1265 GAGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGAGGCGTCCCTCTCG 1324
Db 241 GAGGGGCTTGTTCCTTCCCTCCCGGCGACAAGCTCCAGGAGGCGTCCCTCTCG 300

QY 1325 GCGGCCAGCACTTCTCAGACAACTTCTCTGCTGCTCCAGTCTGGGATCATCA 1384
Db 301 GCGGCCAGCACTTCTCAGACAACTTCTCTGCTGCTCCAGTCTGGGATCATCA 360

QY 1385 CTTACCCACCCGAGTTCAAGACCAATCTTCAGCTGCGCCCTCTGTTTCCCTGT 1444

Db 361 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTGCCCTTCGTGTTCCCTGT 420
QY 1445 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCCCTTGTAACTTCTTAAGTCTAAAGATGATGAACCTTC 1546
Db 481 CTGACCCCTTGTAACTTCTTAAGTCTAAAGATGATGAACCTTC 522

RESULT 7

US-09-759-143-387

; Sequence 387, Application US/09759143

; Patent No. US200202248A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C23

; CURRENT APPLICATION NUMBER: US/09/759,143

; NUMBER OF SEQ ID NOS: 934

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 387

; LENGTH: 537

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-759-143-387

Query Match 33.8%; Score 522; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCCTCTGTGCCATCATGACACCTATGAGTTGCGCAAAAGCTTCTCCAGAGGC 1144
Db 61 CCCCCTCTGTGCCATCATGACACCTATGAGTTGCGCAAAAGCTTCTTCAGAGGC 120
QY 1145 TGAACCAAGACCGGCTTCTTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 1204
Db 121 TGAACCAAGACCGGCTTCTTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 180
QY 1205 CCACGATGGGAGGAGGCGCAGGAGGAGACCCAGCCAAAGTGCCTTTCTCAGACTGAGG 1264
Db 181 CCACGATGGGAGGAGGCGCAGGAGGAGACCCAGCCAAAGTGCCTTTCTCAGACTGAGG 240
QY 1265 GAGGGGGCTTGTTCCTTCCTCCCGGACAGCTCCAGGCGAGGCTGTCCTCTGG 1324
Db 241 GAGGGGGCTTGTTCCTTCCTCCCGGACAGCTCCAGGCGAGGCTGTCCTCTGG 300
QY 1325 GGGGCCCCAGCACTTCCTCAGACAACTTCTTCCTGTGCTCCAGTGTGGGGATCATCA 1384

Db 301 GGGGCCCCAGCACTTCCTCAGACAACTTCTTCCTGTGTCCAGTGTGGGGATCATCA 360
QY 1385 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTGCCCTTCGTGTTCCCTGT 1444
Db 361 CTTACCCACCCCAAGTTCAAGACCAATCTTCAGCTGCCCTTCGTGTTCCCTGT 420
QY 1445 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 1504
Db 421 GTTTCCTGTAGCTGGGATGCTCCAGGAACCAAGAGCCCTCAGCTGTGTAGTCTCC 480
QY 1505 CTGACCCCTTGTAACTTCTTAAGTCTAAAGATGATGAACCTTC 1546
Db 481 CTGACCCCTTGTAACTTCTTAAGTCTAAAGATGATGAACCTTC 522

RESULT 8

US-09-780-669-387

; Sequence 387, Application US/09780669

; Patent No. US20020051977A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C24

; CURRENT APPLICATION NUMBER: US/09/780,669

; NUMBER OF SEQ ID NOS: 943

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 387

; LENGTH: 537

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-780-669-387

Query Match 33.8%; Score 522; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 GGGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 1084
Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTGCAAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1085 CCCCCTCTGTGCCATCATGACACCTATGAGTTGCGCAAAAGCTTCTTCAGAGGC 1144
Db 61 CCCCCTCTGTGCCATCATGACACCTATGAGTTGCGCAAAAGCTTCTTCAGAGGC 120
QY 1145 TGAACCAAGACCGGCTTCTTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 1204
Db 121 TGAACCAAGACCGGCTTCTTGGCGGCTGAAAGGGGCAAGGAGGCAAGCCCGTCTCTC 180
QY 1205 CCACGATGGGAGGAGGCGCAGGAGGAGACCCAGCCAAAGTGCCTTTCTCAGACTGAGG 1264
Db 181 CCACGATGGGAGGAGGCGCAGGAGGAGACCCAGCCAAAGTGCCTTTCTCAGACTGAGG 240

QY 1265 GAGGGGCTGTTTCCCTCCCGCGACAAGCTCCAGGCGAGGGCTGTCCTCTGG 1324
Db 241 GAGGGGCTGTTTCCCTCCCGCGACAAGCTCCAGGCGAGGGCTGTCCTCTGG 300
QY 1325 GCGGCCAGACATCTCTCAGACACAATCTTCTGCTGCTCAGTCGTCGGGATCATCA 1384
Db 301 GCGGCCAGACATCTCTCAGACACAATCTTCTGCTGCTCAGTCGTCGGGATCATCA 360
QY 1385 CTTACCCAGCCCAAGTTCAAGCAAAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 1444
Db 361 CTTACCCAGCCCAAGTTCAAGCAAAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 420
QY 1445 GTTTCCTAGCTGGGATGTTCTCCAGGAACCAAGAGCCCTCAGCCTGTTAGTCTCC 1504
Db 421 GTTTCCTAGCTGGGATGTTCTCCAGGAACCAAGAGCCCTCAGCCTGTTAGTCTCC 480
QY 1505 CTGACCCCTGTTAATCTTAAAGTCTAAAGATGATGAATTC 1546
Db 481 CTGACCCCTGTTAATCTTAAAGTCTAAAGATGATGAATTC 522
RE...
US-09-822-827-387
; Sequence 387, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-387
Query Match 33.8%; Score 522; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1025 GGGCGAGTCGGGACCAAGAGGACTTTTGGAGGCTTCTTCCGATCATCAAGGCTG 1084
Db 1 GGGCGAGTCGGGACCAAGAGGACTTTTGGAGGCTTCTTCCGATCATCAAGGCTG 60
QY 1085 CCCCCTCTGTGCTATCATGATCAGACCTATGATGTTCCGCAAAAGCTTCTTCCAGAGGC 1144
Db 61 CCCCCTCTGTGCTATCATGATCAGACCTATGATGTTCCGCAAAAGCTTCTTCCAGAGGC 120
QY 1145 TGAACCAAGGACCGCTTCTGGGCGCTGAAGGGGCAAGGAGCAAGACCGCTCTC 1204
Db 121 TGAACCAAGGACCGCTTCTGGGCGCTGAAGGGGCAAGGAGCAAGACCGCTCTC 180
QY 1205 CCACGATGGGAGAGGGGAGGAGGAGACCCAGCAAGTGCCTTTTCTCAGCACTGAGG 1264
Db 181 CCACGATGGGAGAGGGGAGGAGGAGACCCAGCAAGTGCCTTTTCTCAGCACTGAGG 240
QY 1265 GAGGGGGCTGTTTCCCTTCCCTCCCGGCGACAGCTCCAGGCGAGGGCTGTCCTCTGG 1324
Db 241 GAGGGGGCTGTTTCCCTTCCCTCCCGGCGACAAAGCTCCAGGCGAGGGCTGTCCTCTGG 300
QY 1325 GGGGCCAGACATCTCTCAGACACAATCTTCTGCTGCTCAGTCGTCGGGATCATCA 1384
Db 301 GGGGCCAGACATCTCTCAGACACAATCTTCTGCTGCTCAGTCGTCGGGATCATCA 360
QY 1385 CTTACCCAGCCCAAGTTCAAGCAAAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 1444
Db 361 CTTACCCAGCCCAAGTTCAAGCAAAATCTTCCAGCTGCCCTTCTGTTTCCCTGT 420
QY 1445 GTTTCCTAGCTGGGATGTTCTCCAGGAACCAAGAGCCCTCAGCCTGTTAGTCTCC 1504

Db 421 GTTTCCTAGCTGGGATGTTCTCCAGGAACCAAGAGCCCTCAGCCTGTTAGTCTCC 480
QY 1505 CTGACCCCTGTTAATCTTAAAGTCTAAAGATGATGAATTC 1546
Db 481 CTGACCCCTGTTAATCTTAAAGTCTAAAGATGATGAATTC 522
RESULT 10
US-09-880-107-1133/C
; Sequence 1133, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1133
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA450247
US-09-880-107-1133
Query Match 28.8%; Score 445.4; DB 10; Length 452;
Best Local Similarity 99.8%; Pred. No. 1.8e-110;
Matches 446; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1100 TCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGC 1159
Db 452 TCATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGC 393
QY 1160 TTCTGGGCGCTGAAGGGGCAAGGAGGCAAGGACCGCTCTCTCCACGGATGGGAGA 1219
Db 392 TTCTGGGCGCTGAAGGGGCAAGGAGGCAAGGACCGCTCTCTCCACGGATGGGAGA 333
QY 1220 GGGCAGGAGGAGACCCAGCAAGTGCCTTTTCTCAGCACTGAGGGAGGGGCTTGTTC 1279
Db 332 GGGCAGGAGGAGACCCAGCAAGTGCCTTTTCTCAGCACTGAGGGAGGGGCTTGTTC 273
QY 1280 CCTTCCCTCCCGGCGACAGCTCCAGGCGAGGGCTGTCCTCTGGGCGGCGCAGCACTTC 1339
Db 272 CCTTCCCTCCCGGCGACAGCTCCAGGCGAGGGCTGTCCTCTGGGCGGCGCAGCACTTC 213
QY 1340 CTCAGACAAACTTCTTCTGCTGCTCCAGTCTGTTGGGATCATCACTTACCCACCCCA 1399
Db 212 CTCAGACAAACTTCTTCTGCTGCTCCAGTCTGTTGGGATCATCACTTACCCACCCCA 153
QY 1400 AGTTCAAGACCAATCTTCCAGCTGCCCTTCTGTTTCCCTGTTGTTGCTGTAGCTGG 1459
Db 152 AGTTCAAGACCAATCTTCCAGCTGCCCTTCTGTTTCCCTGTTGTTGCTGTAGCTGG 93
QY 1460 GCATGCTCCAGGAACCAAGAGGCGCTCAGCTGTTAGTCTCCCTGAGCCCTTGTAAAT 1519
Db 92 GCATGCTCCAGGAACCAAGAGGCGCTCAGCTGTTAGTCTCCCTGAGCCCTTGTAAAT 33
QY 1520 TCCCTAAGTCTAAAGATGATGAATTC 1546
Db 32 TCCCTAAGTCTAAAGATGATGAATTC 6
RESULT 11

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 19:06:17 ; Search time 2432.33 Seconds
(without alignments)
10293.906 Million cell updates/sec

Title: US-09-888-358-2

Perfect score: 1546

Sequence: 1 ggctaggctgcgtgcgagcgcg.....gtctaaagatgaacttc 1546

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Tc number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_esti:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1381.2	89.3	1402	11 AF119864	AF119864 Homo sapi
2	909.6	58.8	1020	13 BM554056	BM554056 AGENCOURT
3	908.8	58.8	963	9 AL519986	AL519986 AL519986
4	888.8	57.5	981	14 BM923995	BM923995 AGENCOURT
5	878.8	56.8	1001	9 AU519987	AL519987 AU519987
6	877.4	56.8	921	9 AU529906	AL529906 AL529906

7	869.4	56.2	920	14 BQ683596	BQ683596 AGENCOURT
8	869	56.2	932	14 BQ225174	BQ225174 AGENCOURT
9	868.4	56.2	925	9 AL516722	AL516722 AL516722
10	867	56.1	911	9 AL528266	AL528266 AL528266
11	864	55.9	910	9 AL570513	AL570513 AL570513
12	855.8	55.4	959	14 BM915493	BM915493 AGENCOURT
13	854	55.2	939	14 BQ949458	BQ949458 AGENCOURT
14	853.8	55.2	965	14 BQ934841	BQ934841 AGENCOURT
15	847	54.8	885	9 AL514846	AL514846 AL514846
16	839.6	54.3	967	13 BI759919	BI759919 603044343
17	838.8	54.3	1077	14 BM808978	BM808978 AGENCOURT
18	837.2	54.2	884	9 AL572559	AL572559 AL572559
19	836.6	54.1	901	9 AL545150	AL545150 AL545150
20	828	53.6	1120	13 BM464148	BM464148 AGENCOURT
21	826.2	53.4	1450	11 AK007934	AK007934 Mus muscu
22	823	53.2	1080	14 BM917004	BM917004 AGENCOURT
23	820.4	53.1	891	9 AL522522	AL522522 AL522522
24	820.2	53.1	890	9 AL515130	AL515130 AL515130
25	818.6	52.9	994	13 BM555499	BM555499 AGENCOURT
26	818	52.9	1012	13 BM468665	BM468665 AGENCOURT
27	813.4	52.6	962	14 BQ688847	BQ688847 AGENCOURT
28	813.4	52.6	1112	14 BM810266	BM810266 AGENCOURT
29	812	52.5	917	9 AL544360	AL544360 AL544360
30	811.4	52.5	1371	11 AK007276	AK007276 Mus muscu
31	809	52.3	872	14 BQ883416	BQ883416 AGENCOURT
32	805.8	52.1	1103	14 BM924115	BM924115 AGENCOURT
33	805	52.1	911	12 BG337246	BG337246 602434762
34	802	51.9	816	12 BG470240	BG470240 602533758
35	801.6	51.8	1453	11 AK002381	AK002381 Mus muscu
36	797.2	51.6	1004	14 BM928609	BM928609 AGENCOURT
37	795.8	51.5	940	9 AL523298	AL523298 AL523298
38	790.2	51.1	847	13 BI333101	BI333101 602996377
39	789.4	51.1	862	12 BG470248	BG470248 602533766
40	788.8	51.0	829	9 AL563104	AL563104 AL563104
41	788.8	51.0	1092	13 BM460198	BM460198 AGENCOURT
42	787.2	50.9	1491	11 AK019396	AK019396 Mus muscu
43	787	50.9	1029	13 BM459839	BM459839 AGENCOURT
44	786.4	50.9	918	12 BG574617	BG574617 602596614
45	785	50.8	795	9 AL571090	AL571090 AL571090

ALIGNMENTS

RESULT 1
AF119864
LOCUS Homo sapiens PRO2163 mRNA, complete cds.
DEFINITION AF119864
ACCESSION AF119864
VERSION AF119864.1 GI:7770164
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
and He,F.
TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1402)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
and He,F.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology,
Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing,
Beijing 100850, P. R. China
FEATURES
source
1. .1402
/location/Qualifiers
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="FLB8041"

FEATURES

Location/Qualifiers

1. 1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5742317"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 175 a 327 c 302 g 213 t 3 others

ORIGIN

Y Match 58.8%; Score 909.6; DB 13; Length 1020;
Local Similarity 96.2%; Pred. No. 8.9e-213;
Matches 982; Conservative 0; Mismatches 12; Indels 27; Gaps 4;

QY 57 GCCTGGCCCTGGCGCGCGCCGACGACACCAAGCCTAGAGCAGGACTGAAGCTTCAA 116
DB 1 GCCTGACCTGGCGCGCGCGCCGACGACACCAAGCCTAGAGCAGGACTGAAGCTTCAA 60
QY 117 GATGCTGACGAGACCTGCGGGGATAGCCCTTCAGCAATGGTGGCTCAGGCAC 176
DB 61 GATGCTGACGAGACCTGCGGGGATAGCCCTTCAGCAATGGTGGCTCAGGCAC 120
QY 177 CGGGCTGTGTACTCTCTTATGACACCCCTGAGTGTGTGAAGTTGCGCTGCA 236
DB 121 CGGGCTGTGTACTCTCTTATGACACCCCTGAGTGTGTGAAGTTGCGCTGCA 180
QY 237 GFTCAGCGCCCTCCATGCGCAGGAGTGTGCTTCTCCAGACTGTGAGGCTCTC 296
DB 181 GFTCAGCGCCCTCCATGCGCAGGAGTGTGCTTCTCCAGACTGTGAGGCTCTC 240
QY 297 CTATACCAAT-----GGAAGTGCCTCTGTATTGCAATGG 332
DB 241 CTATACCAATGTGCTCTCTCTCCAAATCCACAGGAGTGTGCTCTGTATTGCAATGG 300
QY 333 TGCTCTGAGGCTGTACTGTGCCCCAAATGTGCGCGTGTGCCACCTGTGTTTCAGA 392
DB 301 TGCTCTGAGGCTGTACTGTGCCCCAAATGTGCGCGTGTGCCACCTGTGTTTCAGA 360
C 393 CCTACCGCTTACTGGCACCATTGATGCTGTGGAAGATCGTGAGGCACGAGGGCAC 452
DB 361 CCTACCGCTTACTGGCACCATTGATGCTGTGGAAGATCGTGAGGCACGAGGGCAC 420
QY 453 CAGGACCTCTGAGCGGCTCCCGCCACCTGTGTGATGACTGTGCGAGCTACCGCAT 512
DB 421 CAGGACCTCTGAGCGGCTCCCGCCACCTGTGTGATGACTGTGCGAGCTACCGCAT 480
QY 513 CTACTTCACTGCTATGACCAATGAAGGCTTCTGTGTGTGCGAGCCTGACCTCTGA 572
DB 481 CTACTTCACTGCTATGACCAATGAAGGCTTCTGTGTGTGCGAGCCTGACCTCTGA 540
QY 573 CTTTACGCAACCATGTGTGGCGGCTGGCCGCTGGCCGCTGGCCACCTGACTGTGATCAG 632
DB 541 CTTTACGCAACCATGTGTGGCGGCTGGCCGCTGGCCGCTGGCCACCTGACTGTGATCAG 600
QY 633 CCCCTGAGCTTATGCGCAAAAGTCAGGCTCAGCATGTGTGTACCGGAGCTGGG 692
DB 601 CCCCTGAGCTTATGCGCAAAAGTCAGGCTCAGCATGTGTGTACCGGAGCTGGG 660
QY 693 TGCCTGTGTGCAACTCAGTGTGCTCAGGGTGGCTGGGCTCACTGTGCTGGGCTGGG 752
DB 661 TGCCTGTGTGCAACTCAGTGTGCTCAGGGTGGCTGGGCTCACTGTGCTGGGCTGGG 720
QY 753 CCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACT-GGTTCAACTATGAGCTGG 811

DB 721 CCCACTGCCCTTCAGATGTGCCCTTCTCAGCCCTGTACTGGTTCACATGAGCTGG 780
QY 812 TGAAGAGCTGGCTCAATGGTTTCCAGCCGAGAGCAGACCTTCTGTGGGATGAGCTTTG 871
DB 781 TGAAGAGCTGGCTCAATGGTTTCCAGCCGAGAGCAGACCTTCTGTGGGATGAGCTTTG 840
QY 872 TGGCTGGTGGCATCTCAGGAGCGGTGGTGCAGTCTGACTCTACCCCTTTGAGCTGTAA 931
DB 841 TGGCTGGTGGCATCTCAGGAGCGGTGGTGCAGTCTGACTCTACCCCTTTGAGCTGTAA 899
QY 932 AGACCAACGCGCAGGTGGCTCT-GGAGCGATGAGGCTGTGAGAGTGAACCCCTGCAAT 990
DB 900 AGACCAACGCGCAGGTGGCTCTGCGGAAAGTGTGANAATGAACCCCTGCAAT 959
QY 991 GTGAGTCACTGCTGGTGTGCTGCGAGGATCCGGCCGAGTCCGGCACCAGGAGCTC 1050
DB 960 GTGAGTCACTGCTGGTGTGCTGCGAGGATCCGGCCGAGTCCGGCACCAGGAGCTC 1019
QY 1051 T 1051
DB 1020 T 1020

RESULT 3
AL519986/c
LOCUS AL519986 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB005YF09 3
DEFINITION prime, mRNA sequence.
ACCESSION AL519986
VERSION AL519986.1 GI:12783479
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB005YF09"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 206 a 288 c 298 g 168 t 3 others

ORIGIN

Query Match 58.8%; Score 908.8; DB 9; Length 963;
Best Local Similarity 98.2%; Pred. No. 1.4e-212;
Matches 947; Conservative 3; Mismatches 11; Indels 3; Gaps 3;

QY 563 TGACCTCTGACCTCTACGACCCATGTTGGTGGCGGCTGGCCGCTGGGACCGCTGA 622
DB 963 TGACCTCTGACCTCTACGACCCATGTTGGT-TKGGCGCTGGCCGCTGGGACCGCTGA 905

QY	623	CTGTGATCAGCCCCCTTGAGCTTATCGGACAAAGCTGCAGGCTCAGCACTGTGTCTGTACC	682
DB	904	CTGTGATCAGCCCCCTTGAGCTTATCGGACAAAGCTGCAGGCTCAGCATGTGTCTGTACC	845
QY	683	GGGAGCTGGGTGCTGTGTCGAACTGCAGTGGCTCAGGGTGGCTGGCGCTCACTGTGGC	742
DB	844	GGGAGCTGGGTGCTGTGTCGAACTGCAGTGGCTCAGGGTGGCTGGCGCTCACTGTGGC	785
QY	743	TGGGCTGGGCCCCACTGCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTCAACT	802
DB	784	TGGGCT - GGGCCCCACTGCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTCAACT	726
QY	803	ATGAGCTGGTGAAGAGCTGGCTCAATGGCTTCAGGCCGAGGAGCAGAGACTTCGTGGGCA	862
DB	725	ATGAGCTGGTGAAGAGCTGGCTCAATGGCTTCAGGCCGAGGAGCAGAGACTTCGTGGGCA	666
QY	863	TGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGCAGTGTGACTTACCCCTTTG	922
DB	665	TGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGCAGTGTGACTTACCCCTTTG	606
QY	923	ACGTGTTAAAGACCCCAAGCCAGGTGCCTCTGGAGAGCATGAGAGCTGTGAGAGTAGAAC	982
DB	605	ACGTGTTAAAGACCCCAAGCCAGGTGCCTCTGGAGAGCATGAGAGCTGTGAGAGTAGAAC	546
QY	983	CCCTGCATGTGACTCCACTGTGCTGCTGCGAGAGATCGGGCCGAGTCGGGCAACA	1042
DB	545	CCCTGCATGTGACTCCACTGTGCTGCTGCGAGAGATCGGGCCGAGTCGGGCAACA	486
QY	1043	AGGAGCTCTTTGAGGCTTCCTTCCTGGATCATCAAGGCTGCCCTTCCTGTGCCATCA	1102
DB	485	AGGAGCTCTTTGAGGCTTCCTTCCTGGATCATCAAGGCTGCCCTTCCTGTGCCATCA	426
QY	1103	TGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGCTGAACAGACCGGCTTC	1162
DB	425	TGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGCTGAACAGACCGGCTTC	366
QY	1163	TGGCGGCTGAAAGGGGCAAGGAGGCAAGGACCCCGCTCTCCACGGATGGGAGAGGG	1222
DB	365	TGGCGGCTGAAAGGGGCAAGGAGGCAAGGACCCCGCTCTCCACGGATGGGAGAGGG	306
QY	1223	CAGGAGAGACCCAGCAAGTGTCTTTCTCAGCATGAGGAGGGGGCTTGTTCCT	1282
DB	305	CAGGAGAGACCCAGCAAGTGTCTTTCTCAGCATGAGGAGGGGGCTTGTTCCT	246
QY	1283	TCCCTCCCGGACAAAGCTCCAGGCGAGGGTGTCCCTCTGGGCGGCCAGACACTTCCT	1342
DB	245	TCCCTCCCGGACAAAGCTCCAGGCGAGGGTGTCCCTCTGGGCGGCCAGACACTTCCT	186
QY	1343	AGACACAACCTTCTCTGCTGCTCCAGTGGTGGGATCATCACTTACCCACCCCCCAAGT	1402
DB	185	AGACACAACCTTCTCTGCTGCTCCAGTGGTGGGATCATCACTTACCCACCCCCCAAGT	126
QY	1403	TCAAGACCAATCTTCAGAGTGGCCCTTCGTGTGTTTCCCTGTGTTTGTGTAGCTGGGCA	1462
DB	125	TCAAGACCAATCTTCAGAGTGGCCCTTCGTGTGTTTCCCTGTGTTTGTGTAGCTGGGCA	67
QY	1463	TGCTCTCAGGAACCAAGAAGCCCTCAGCCTGGTGTAGTCTCCCTCAGCCCTTGTTAAATCC	1522
DB	66	TGCTCTCAGGAACCAAGAAGCCCTCAGCCTGGTGTAGTCTCCCTCAGCCCTTGTTAAACCT	7
QY	1523	TTAA	1526
DB	6	TAAD	3

```

RESULT 4
BM923995
LOCUS
DEFINITION
      BM923995 981 bp mRNA linear EST 12-MAR-2002
      AGENCOURT_6709842 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760154
      5', mRNA sequence.
ACCESSION
      BM923995
      BM923995.1 GI:19374374
VERSION

```

```

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 981)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL Unpublished (1999)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12806 Row: 1 Column: 11
High quality sequence stop: 701.
FEATURES
Location/Qualifiers
1..981
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5760154"
/clone_lib="NIH MGC_116"
/lab_host="DH10B"
/path="/Orgen: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site_1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."
BASE COUNT 171 a 309 c 298 g 201 t
ORIGIN

```

Query Match	57.5%; Score 888.8; DB 14; Length 981;
Best Local Similarity	97.2%; Pred. No. 1.1e-207;
Matches 946; Conservative 0; Mismatches 27; Indels 4; Gaps 4;	
QY 33	AGGGGCGACGAGCTGAATGGGCGCGCTGGCGCTGGGCGCGCGCGACGAGCACGCG 92
Db 5	AGGCGGGGAGCGGCTATGGCGCGCGCTGGCGCTGGGCGCGCGCGACGAGCACGCG 64
QY 93	CTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACACGAGCCCTCGGGGCATCAGCCCCCT 152
Db 65	CTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACACGAGCCCTCGGGGCATCAGCCCCCT 124
QY 153	CCAGCAAAATGTCGGCTCAGGACACGGGGCTGTGGTTACTCTCTCTCTGACACCCCT 212
Db 125	CCAGCAAAATGTGGCTCAGGACACGGGGCTGTGGTTACTCTCTCTCTGACACCCCT 184
QY 213	GGACGTGTTGAAGGTTTCGCTGCAGTCTCAGCGGGCCCTCCATGGCAGCGAGCTGATGCC 272
Db 185	GGACGTGTTGAAGGTTTCGCTGCAGTCTCAGCGGGCCCTCCATGGCAGCGAGCTGATGCC 244
QY 273	TTCCTCCAGAGCTGGAGGCGCTCCCTATACAAATGGAAGTGCCTCTGTATGCAATGG 332
Db 245	TTCCTCCAGAGCTGGAGGCGCTCCCTATACAAATGGAAGTGCCTCTGTATGCAATGG 304
QY 333	TGTCTCTGGAGCCTCTGTACCTGTGCCCAAAATGGTGGCCCGCTGTGCCACCTGGTTCAAGA 392
Db 305	TGTCTCTGGAGCCTCTGTACCTGTGCCCAAAATGGTGGCCCGCTGTGCCACCTGGTTCAAGA 364
QY 393	CCCTACCCGGCTTCACTGGCACCATGGATGCCTTCGTGAAGATCGTGAGGCACGAGGGCAC 452
Db 365	CCCTACCCGGCTTCACTGGCACCATGGATGCCTTCGTGAAGATCGTGAGGCACGAGGGCAC 424

QY 453 CAGGACCTCTGGAGCGGCTCCGCCACCTGGTGTGATGCTGTGCGAGCTACCGCAT 512
 Db 425 CAGGACCTCTGGAGCGGCTCCGCCACCTGGTGTGATGCTGTGCGAGCTACCGCAT 484
 QY 513 CTACTTCACTGCTATGACCAACTGAAGCCCTTCTGTGTGCTGAGCCCTGACCTTGA 572
 Db 485 CTACTTCACTGCTATGACCAACTGAAGCCCTTCTGTGTGCTGAGCCCTGACCTTGA 544
 QY 573 CTTCTACGACCCATGCTGTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 632
 Db 545 CTTCTACGACCCATGCTGTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 604
 QY 633 CCCCCTGGAGCTTATGCGGCAAAAGCTGACAGCTCAGCATGTGCTGACCGGAGCTGGG 692
 Db 605 CCCCCTGGAGCTTATGCGGCAAAAGCTGACAGCTCAGCATGTGCTGACCGGAGCTGGG 664
 QY 693 TGCCTGTGTTCAAACTGCACTGAGCTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 752
 Db 665 TGCCTGTGTTCAAACTGCACTGAGCTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 724
 QY 753 CCCCCTGGAGCTTATGCGGCAAAAGCTGACAGCTCAGCATGTGCTGACCGGAGCTGGG 812
 Db 725 CCCCCTGGAGCTTATGCGGCAAAAGCTGACAGCTCAGCATGTGCTGACCGGAGCTGGG 784
 QY 813 GAAGAGCTGGCTCAATGGGTTCAAGCGG-C-AAGGACCAAGCTTCTGTGGGCAATGAGC-TTT 870
 Db 785 GAAGAGCTGGCTCAATGGGTTCAAGCGG-C-AAGGACCAAGCTTCTGTGGGCAATGAGC-TTT 844
 QY 871 GTGGCTGTGGGATCTCAGGAGCGGTGGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 929
 Db 845 GTGGCTGTGGGATCTCAGGAGCGGTGGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 904
 QY 930 AAGAGCCCAAGCCAGCTGCTCTGGGAGGAT-CGAGGCTGAGAGTGAACCCCTGCG 988
 Db 905 AAGAGCCCAAGCCAGCTGCTCTGGGAGGAT-CGAGGCTGAGAGTGAACCCCTGCG 964
 QY 989 ATGTGAGCTCCAC 1001
 Db 965 ATGGGAGCTCCAC 977
 RESULT 5
 AL519987
 LOCUS
 DEFINITION AL519987 LTI_NFL004_NEC2 Homo sapiens cDNA clone EST 13-FEB-2001
 prime, mRNA sequence.
 ACCESSION AL519987
 VERSION AL519987.1 GI:12783480
 X RDS
 S _E human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1. (bases 1 to 1001)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 FEATURES
 source
 1. .1001
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSDB005YP09"
 /lib_host="LTI_NFL004_NEC2"
 /sex="male"
 /tissue type="neuroblastoma cells"
 /lab_host="PH10B"
 /notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 167 a 309 c 315 g 206 t 2 others
 ORIGIN
 Query Match 56.8%; Score 878.8; DB 9; Length 1001;
 Best Local Similarity 97.6%; Pred. No. 3.3e-205;
 Matches 955; Conservative 0; Mismatches 17; Indels 6; Gaps 6;
 QY 28 CCACGAGGCGGACGGAGCTAATGGGCGGCTTGGCCCTGGCCCTGGGCGCGCCGCGCACGAGCA 87
 Db 11 CCACGAGGCGGCGGAGCTAATGGGCGGCTTGGCCCTGGCCCTGGGCGCGCCGCGCACGAGCA 70
 QY 88 CCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACAGGACCCCTGCGGCATCAGC 147
 Db 71 CCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACAGGACCCCTGCGGCATCAGC 130
 QY 148 CCCCCTCAGCAAAATGTGGGCTCAGGCA CCGGGCTGTGGTTACTCTCTCTTTTCATGACA 207
 Db 131 CCCCCTCAGCAAAATGTGGGCTCAGGCA CCGGGCTGTGGTTACTCTCTCTTTTCATGACA 190
 QY 208 CCCCCTGAGCTGTGAGGTTGCGCTTGCAGTCTCAGCGGCCCTCCATGGCCACGAGCTG 267
 Db 191 CCCCCTGAGCTGTGAGGTTGCGCTTGCAGTCTCAGCGGCCCTCCATGGCCACGAGCTG 250
 QY 268 ATGCTTCTCCAGACTGTGGAGCTCTCTATACAAATGGAAGTCCCTCTGTATATTC 327
 Db 251 ATGCTTCTCCAGACTGTGGAGCTCTCTATACAAATGGAAGTCCCTCTGTATATTC 310
 QY 328 AATGCTGCTTGGAGCTCTGTACTGTGCCCCAAATGTGCCCCCTGTGCCACCTGGTTT 387
 Db 311 AATGCTGCTTGGAGCTCTGTACTGTGCCCCAAATGTGCCCCCTGTGCCACCTGGTTT 370
 QY 388 CAAGACCTTACCGCTTCACTGGGACCATGGATGCTTCTGTGAAGATCGTGAGGACAGAG 447
 Db 371 CAAGACCTTACCGCTTCACTGGGACCATGGATGCTTCTGTGAAGATCGTGAGGACAGAG 430
 QY 448 GGACACGAGACCTCTGGAGCGGCTCCCGCCACCTGGTGATGACTGTGCCAGCTACC 507
 Db 431 GGACACGAGACCTCTGGAGCGGCTCCCGCCACCTGGTGATGACTGTGCCAGCTACC 490
 QY 508 GCCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGGTGGAGCCCTGACC 567
 Db 491 GCCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGGTGGAGCCCTGACC 550
 QY 568 TCTGACCTTACGACCCATGGTGGCTGGGCGGCTGGCCCGCTGGGACACGCTGACTGTG 627
 Db 551 TCTGACCTTACGACCCATGGTGGCTGGGCGGCTGGCCCGCTGGGACACGCTGACTGTG 610
 QY 628 ATCAGCCCCCTGGAGCTTATGCGGCAAAAGCTGACAGGCTCAGCATGTGCTACCGGGAG 687
 Db 611 ATCAG-CCCTGGAGCTTATGCGGCAAAAGCTGACAGGCTCAGCATGTGCTACCGGGAG 669
 QY 688 CTGGGTGCTGTGTGGAACCTGCACTGAGTGGCTGAGGCTGGCTGGCTCACTGTGGCTGGGC 747
 Db 670 CTGGGTGCTGTGTGGAACCTGCACTGAGTGGCTGAGGCTGGCTGGCTCACTGTGGCTGGGC 729
 QY 748 TGGGGCCCCACTGCCCCCTCAGATGTGCCCTTCTCAGCCCTGTACTGGTCAACTATGAG 807
 Db 730 TGGGG-CCCACTGCCCCCTCAGATGTGCCCTTCTCAGCCCTGTACTGGTCAACTATGAG 788
 QY 808 CTGCTGAAGAGCTGGCTCAATGGGCTCAGGCCGAA- GACACAGACTTCTGTGGGATGAGC 867
 Db 789 CTGCTGAAGAGCTGGCTCAATGGGCTCAGGCCGAA- GACACAGACTTCTGTGGGATGAGC 847
 QY 968 TTTGTGGCTGTGGCATCTCAGGAGCGGTGGCTGCACTGTGAGTCTTACCCCTTTGACGTTG 927
 Db 848 TTTGTGG- TGTGTGTCATCTCAGGAGCGGTGGCTGCACTGTGAGTCTTACCCCTTTGACGTTG 907

```

QY 928 GTAAAGACCCAAACCCAGGTCGCTCTGGAGCGATGAGAGCTGTGAGATGAACCCCTG 987
Db 908 GTAAAGA-CCAAACCCAGGTCGCTCTGGAG-GATGGAGGTGTGAGATGA-CCCTG 964
QY 988 CATGTGGACTCCACTGG 1005
Db 965 CATGTGGACTCCACTGG 982

RESULT 6
AL529906 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD005YL19 5
LOCUS DEFINITION prime, mRNA sequence.
ACCESSION AL529906
VERSION AL529906.1 GI:12793399
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 921)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..921
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DD005YL19"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Li
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 154 a 292 c 269 g 193 t 13 others
ORIGIN

Query Match 56.8%; Score 877.4; DB 9; Length 921;
Best Local Similarity 97.5%; Pred. No. 7e-205;
Matches 889; Conservative 10; Mismatches 12; Indels 1; Gaps 1;

QY 35 GCGGACGCGACGTAAATGGGCGCCCTGGCCCTGGCGCGCGCGGCGGACGACGACGCT 94
Db 11 GCGGCGCGGASGNTATTGGGCGCCCTGGCCCTGGCGCGCGCGGCGGACGACGACGCT 70
QY 95 AGAGCAGGACTGAAGCTTCAAGATGGCTGACGAGCCCTGGCGGCGGCGGCGGCGGCTCC 154
Db 71 AGAGCAGGACTGAAGCTTCAAGATGGCTGACGAGCCCTGGCGGCGGCGGCGGCGGCTCC 130
QY 155 AGCAATGGTGGCTCAGGACCGCGGCGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGG 214
Db 131 AGCAATGGTGGCTCAGGACCGCGGCGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGG 190
QY 215 ACCTGGTGAAGTTGGCTCGAGTCTCAGGCGCCCTCATGCGGCGGCGGCGGCTGTGGCTT 274
Db 191 ACCTGGTGAAGTTGGCTCGAGTCTCAGGCGCCCTCATGCGGCGGCGGCGGCTGTGGCTT 250
QY 275 CCTCAGGACTGTGAGGCTCTCTATACCAATGGAAGTCTCTCTATGCAATGGTG 334

```

```

Db 251 CTTCCAGACTGTGAGGCTCTCTATACCAATGGAAGTCTCTGTATTCGAATGGTG 310
QY 335 TCTTGAGACCTCTGTACTGTGCCCCAAATGGTGGCCGCTGTGCCACCTGTGTTTCAAGACC 394
Db 311 TYCTGAGGCTCTGTACTGTGCCCCAAATKGTGGCCGCTGTGCCACCTGTGTTTCAAGACC 370
QY 395 CTACCCGTTTCACTGGCACCATGATGCTTCTGTGAAGATCGTGAGGACAGAGGCACCA 454
Db 371 CTA/CCGTTTCACTGGCACCATGATGCTTCTGTGAAGATCGTGAGGACAGAGGCACCA 430
QY 455 GCACCTCTGAGGCGCTCCCGCACCCCTGTGTGATGACTGTGCCAGCTACGCCATCT 514
Db 431 GCACCTCTGAGGCGCTCCCGCACCCCTGTGTGATGACTGTGCCAGCTACGCCATCT 490
QY 515 ACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTGAGGCGCTGACCTCTGACC 574
Db 491 ACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGTGAGGCGCTGACCTCTGACC 550
QY 575 TCTAGCACCATGTTGGCTGGGCGGCTGGCCGCTGGGACCCGCTGACTGTGATCAGCC 634
Db 551 TCTAGCACCATGTTGGCTGGGCGGCTGGCCGCTGGGACCCGCTGACTGTGATCAGCC 610
QY 635 CCTCGGACTTATCGGACAAAGCTGACAGGCTCAGCATGTGTCGTACCGGAGCTGGTG 694
Db 611 CCTCGGACTTATCGGACAAAGCTGACAGGCTCAGCATGTGTCGTACCGGAGCTGGTG 670
QY 695 CTTGTGTTGMACTGCAAGTGGCTGAGGTTGGGCTGGCCCTCACTGTGGCTGGCTGGGCC 754
Db 671 CTTGTGTTGMACTGCAAGTGGCTGAGGTTGGGCTGGCCCTCACTGTGGCTGGCTGGGCC 730
QY 755 CCACCTGCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAACTATGAGCTGGTGA 814
Db 731 CCACCTGCTTCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAACTATGAGCTGGTGA 790
QY 815 AGAGTGGCTCAATGGGTTTCAGGCGGAGGACACAGACTTCTGTGGGCGATGAGCTTGTGG 874
Db 791 AGAGTGGCTCAATGGGTTTCAGGCGGAGGACACAGACTTCTGTGGGCGATGAGCTTGTGG 849
QY 875 CTGTGGGATCTCAGGACCGGTGGCTGAGTGTGCTGACTCTACCTTTTGGCTGGTAAAGA 934
Db 850 CTGTGGGATCTCAGGACCGGTGGCTGAGTGTGCTGACTCTACCTTTTGGCTGGTAAARA 909
QY 935 CCCAAGCCAGG 946
Db 910 CCCAAGCCAGG 921

RESULT 7
BO683596
LOCUS DEFINITION 920 bp mRNA linear EST 15-JUL-2002
AGENCOURT_8059897 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091198
5', mRNA sequence.
ACCESSION BO683596
VERSION BO683596.1 GI:21796275
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

Plate: LLCM2333 row: m column: 23
High quality sequence stop: 697.

FEATURES

```

source
1..920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6091198"
/clone_lib="NIH MGC 112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT      161 a 288 c 273 g 196 t      2 others
OP: ^N

%-xy Match      56.2%; Score 869.4; DB 14; Length 920;
Best Local Similarity 99.7%; Pred. No. 6.5e-203;
Matches 881; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 75 CGCCGACGAGCACCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACGAGACCC 134
Db 1 CGCCGACAGCAGCAGCCTAGAGCCAGGACTGAAGCTTCAAGATGGCTGACGAGACCC 60

QY 135 TCGGGGATGACGCCCTCCAGCAATGTGGCTCAGGACCGGGCTGTGGTTACCTC 194
Db 61 TCGGGGATGACGCCCTCCAGCAATGTGGCTCAGGACCGGGCTGTGGTTACCTC 120

QY 195 TCTCTTCATGACACCCCTGAGCTGTGAAGTTTCGCTCGAGTCTCAGCGCCCTCCAT 254
Db 121 TCTCTTCATGACACCCCTGAGCTGTGAAGTTTCGCTCGAGTCTCAGCGCCCTCCAT 180

QY 255 GSCCAGCGAGCTGATGCTTCTCCAGAGCTGTGGAGCTTCTCTATACCAATGGAAGTG 314
Db 181 GSCCAGCGAGCTGATGCTTCTCCAGAGCTGTGGAGCTTCTCTATACCAATGGAAGTG 240

QY 315 CTCTCTGTATTGCAATGGTGTCTGAGGCTCTGTACTGTGGCCCAATGGTGGCCGGTG 374
Db 241 CTCTCTGTATTGCAATGGTGTCTGAGGCTCTGTACTGTGGCCCAATGGTGGCCGGTG 300

QY 375 TGCCACCTGGTTTCAAGACCTTACCGCTTCACTGGCACCAGTGGCTTCTGTGAAGAT 434
Db 301 TGCCACCTGGTTTCAAGACCTTACCGCTTCACTGGCACCAGTGGCTTCTGTGAAGAT 360

QY 435 CDTGAGGACGAGAGGACACGACCTCTGAGCGGCTTCCCGCCACCCCTGGTGATGAC 494
Db 361 CDTGAGGACGAGGAGGACACGACCTCTGAGCGGCTTCCCGCCACCCCTGGTGATGAC 420

QY 495 TGTGCGACGTACCGGCTCTACTCTGCTATGACCACTGAGGCTTCTGTGTGG 554
Db 421 TGTGCGACGTACCGGCTCTACTCTGCTATGACCACTGAGGCTTCTGTGTGG 480

QY 555 TGGAGCCCTGACCTCTGACCTCTACGACCCATGTGTGGCTGGCGCTGGCCGCTGGG 614
Db 481 TGGAGCCCTGACCTCTGACCTCTACGACCCATGTGTGGCTGGCGCTGGCCGCTGGG 540

QY 615 CACCGTGACTGTGATCAGCCCTTGGAGCTTATGGGCAAAAGCTGACGGCTCAGCATGT 674
Db 541 CACCGTGACTGTGATCAGCCCTTGGAGCTTATGGGCAAAAGCTGACGGCTCAGCATGT 600

QY 675 GTCTGTACGGAGCTGGGTGCTGTGTTGCACTGAGTGGCTCAGGCTGGCGCTC 734
Db 601 GTCTGTACGGAGCTGGGTGCTGTGTTGCACTGAGTGGCTCAGGCTGGCGCTC 660

QY 735 ACTGTGGCTGGGCTGGGGCCCACTGCCCTTCGAGATGCGCTTCTCAGCCCTGTACTG 794
Db 661 ACTGTGGCTGGGCTGGGGCCCACTGCCCTTCGAGATGCGCTTCTCAGCCCTGTACTG 720

```

```

QY 795 GTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGTTTCAAGCCGAGGACGACACTTC 854
Db 721 GTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGTTTCAAGCCGAGGACGACACTTC 780

QY 855 TGTGGGCTATGAGCTTTTGTGGCTGGTGGCATCTCAGGACCGTGGCTGCAGTGTGACTCT 914
Db 781 TGTGGGCTATGAGCTTTTGTGGCTGGTGGCATCTCAGGACCGTGGCTGCAGGCTGACTCT 840

QY 915 ACCCTTTGACGTGTAAAGACCCCAAGCCAGGCTGGCTCTCGGA 957
Db 841 ACCCTTTGACGTGTAAAGACCCCAAGCCAGGCTGGCTCTCGGA 884

RESULT 8
B0925174
LOCUS B0925174
DEFINITION BQ925174.1 GI:22340205
ACCESSION BQ925174
VERSION BQ925174.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: I.M.G.E. clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2600 row: p column: 15
High quality sequence stop: 616.

FEATURES
source
1..932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6421038"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT      154 a 284 c 290 g 201 t      3 others
ORIGIN
Query Match      56.2%; Score 869; DB 14; Length 932;
Best Local Similarity 97.7%; Pred. No. 8.2e-203;
Matches 911; Conservative 0; Mismatches 18; Indels 3; Gaps 3;

QY 538 AAGGCTTCCTGTGTGGTTCAGGCGCTGACCTCTGACCTTACGACCCCATGGTGGCTGGC 597
Db 1 AAGGCTTCCTGTGTGGTTCAGGCGCTGACCTCTGACCTTACGACCCCATGGTGGCTGGC 60

QY 598 GGGCTGGCGCCCTGGGACCGGCTGACTGTGATCAGCCCTCGAGCTTATGCGGACAAAG 657
Db 61 GGGCTGGCGCCCTGGGACCGGCTGACTGTGATCAGCCCTCGAGCTTATGCGGACAAAG 120

QY 658 CTGCGAGGCTCAGCATGTGTCTGTAACGGGAGCTGGGTGCTGTGTTCGAACTGCAGTGGCT 717

```


Db 121 CTGACGCTCAGCATGTCGTTACCGGAGCTGGGTGCTGTGTTCAACTCAGTGGCT 180
 QY 718 CAGGCTGCTGGCTCCTCAGTGTGCTGGGCTGGGCCCCCAGCTGCTGCTGAGATGCTGCC 777
 Db 181 CAGGCTGCTGGCTCCTCAGTGTGCTGGGCTGGGCCCCCAGCTGCTGAGATGCTGCC 240
 QY 778 TTCTCAGCCCTGCTACTGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGTTTCAAG 837
 Db 241 TTCTCAGCCCTGCTACTGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGTTTCAAG 300
 QY 838 CCGAGGACACAGACTCTGTGGGCACTGAGCTTTGTGGCTGGTGGCTGCTCAGGAGCGGTG 897
 Db 301 CCGAGGACACAGACTCTGTGGGCACTGAGCTTTGTGGCTGGTGGCTGCTCAGGAGCGGTG 360
 QY 898 GCTGAGTGTGCTACTTACCTTTGACCTGTTGAGAGAGCCCAAGCCAGGCTGCTCTGGGA 957
 Db 361 GCTGAGTGTGCTACTTACCTTTGACCTGTTGAGAGAGCCCAAGCCAGGCTGCTCTGGGA 420
 QY 958 GCGATGGAGCTGTGAGAGTGAACCCCTGATGTGAGCTTCCAGCTGGCTGCTGCTGGG 1017
 Db 421 GCGATGGAGCTGTGAGAGTGAACCCCTGATGTGAGCTTCCAGCTGGCTGCTGCTGGG 480
 QY 1018 AGGATCCGGGCGGAGTGGGCACTGTTGAGGAGCTTCTTGGAGGCTTCTTCTCGGATCATC 1077
 Db 481 AGGATCCGGGCGGAGTGGGCACTGTTGAGGAGCTTCTTGGAGGCTTCTTCTCGGATCATC 540
 QY 1078 AAGGCTGCCCCCTCTCTGTGCTATCATGATGATGATGATGATGATGATGATGATGATGAT 1137
 Db 541 AAGGCTGCCCCCTCTCTGTGCTATCATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 1138 CAGAGGCTGAACAGGACCGCTTCTGGGCGCTGAAAGGGGCAAGAGGAGGAGGAGGAGGAG 1197
 Db 601 CAGAGGCTGAACAGGACCGCTTCTGGGCGCTGAAAGGGGCAAGAGGAGGAGGAGGAGGAG 660
 QY 1198 GTCTCT-CCACGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1256
 Db 661 GTCTCTCCACGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 QY 1257 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1316
 Db 721 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
 QY 1317 CCCTCTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1375
 Db 781 CCCTCTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 QY 1376 GGATCATCACTTACCCACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1435
 Db 841 GGATCATCACTTACCCACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
 QY 1436 TTTCCTCT-CTTTGCTGTGAGTGGGCACTGTC 1466
 Db 901 TTTCCTCTGTTGCTGTAACTGGGCACTGGC 932

RESULT 9
 AL516722
 LOCUS
 DEFINITION AL516722 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA007YA10 5 prime, mRNA sequence.
 ACCESSION AL516722
 VERSION AL516722.1 GI:12780215
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 925)
 Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL
 CONTACT: Genoscope
 COMMENT: Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. 925
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DA007YA10"
 /clone_lib="LTI_NFL011_NBC1"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 153 a 294 c 279 g 192 t 7 others
 ORIGIN

Query Match 56.2%; Score 868.4; DB 9; Length 925;
 Best Local Similarity 97.6%; Pred. No. 1.1e-202;
 Matches 905; Conservative 6; Mismatches 13; Indels 3; Gaps 3;

QY 33 AGGCGGAGCAGCTAATGAGCGGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 92
 Db 1 AGGCGGAGCAGCTAATGAGCGGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
 QY 93 CTAGAGCCAGGAGCTGAAGCTTCAAGATGGCTGACCCAGGAGCCCTGCGGGCATCAGCCCTT 152
 Db 61 CTAGAGCCAGGAGCTGAAGCTTCAAGATGGCTGACCCAGGAGCCCTGCGGGCATCAGCCCTT 120
 QY 153 CAGCAAAATGGTGGCTCAGGACACCGGGGCTGTGGTTACCTTCTCTTTCATGACACCTT 212
 Db 121 CAGCAAAATGGTGGCTCAGGACACCGGGGCTGTGGTTACCTTCTCTTTCATGACACCTT 180
 QY 213 GGACGTGTGAAGGCTCGCTCGAGTCTCAGGGCGCTCCATGCGCGAGCTGATGCC 272
 Db 181 GGACGTGTGAAGGCTCGCTCGAGTCTCAGGGCGCTCCATGCGCGAGCTGATGCC 240
 QY 273 TTCTCCAGACTGTGAGGCTCTCTATACCAATGGAAGTGGCTCTCTATTGCAATGG 332
 Db 241 TTCTCCAGACTGTGAGGCTCTCTATACCAATGGAAGTGGCTCTCTATTGCAATGG 300
 QY 333 TGTCTGAGGCTCTGTACCTGTGCCCCAAATGGTGGCGCTGTGCGACCTGGTTTCAAGA 392
 Db 301 TGTCTGAGGCTCTGTACCTGTGCCCCAAATGGTGGCGCTGTGCGACCTGGTTTCAAGA 360
 QY 393 CCTACCCGCTTCACTGGCAGCATGATGCTTCTGTAAGATCTGAGGACAGGAGGAC 452
 Db 361 CCTACCCGCTTCACTGGCAGCATGATGCTTCTGTAAGATCTGAGGACAGGAGGAC 420
 QY 453 CAGGACCTCTGAGGCGGCTCCCCCGCACCTTCTGTGATGACTGTGCCAGCTACCGCCAT 512
 Db 421 CAGGACCTCTGAGGCGGCTCCCCCGCACCTTCTGTGATGACTGTGCCAGCTACCGCCAT 480
 QY 513 CTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGGTGGAGGCTTCACTCTGA 572
 Db 481 CTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGGTGGAGGCTTCACTCTGA 540
 QY 573 CCTTACGACCCCATGTGCTGGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 632
 Db 541 CCTTACGACCCCATGTGCTGGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 600
 QY 633 CCCCCTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGTCTGCTACCGGAGCTGG 692
 Db 601 CCCCCTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGTCTGCTACCGGAGCTGG 660

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1. .910
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="CSOD1018YN07"
/clone.lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 197 a 260 c 272 g 156 t 25 others

ORIGIN

Query Match 55.9%; Score 864; DB 9; Length 910;
Best Local Similarity 96.0%; Pred. No. 1.4e-201;
Matches 868; Conservative 24; Mismatches 11; Indels 1; Gaps 1;

QY 539 AGGCTTCTCTGTGTGTGAGCCCTGACCTCTGACCTTACGACCCATGTGTGGTGGG 598
DB 903 ARGCTTCTCTGTGTGTGAGCCCTGACCTTCTGACCTTACGACCCATGTGTGGTGGG 844
QY 599 CGTGTGCGCCCTGGGACCCCTGACTGTGATCAGCCCTGGAGCTTATGGGACAAAGC 658
DB 843 CGCTGTGCGCCCTGGGACCCCTGACTGTGATCAGCCCTGGAGCTTATGGGACAAAGC 784
QY 659 TGAGGCTAGATGTGCTGACGGAGCTGGGTGCTGTGTTGAACTGCTGAGTGGCTC 718
DB 783 TGCAGGCTCAGCATGTGCTGACGGAGCTGGGTGCTGTGTTGAACTGCTGAGTGGCTC 724
QY 719 AGGTGCTGGGCTGCTGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 778
DB 723 AGGTGCTGGGCTGCTGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 665
QY 779 TCTCAGCCCTGTACTGCTTCAACTATGAGTGTGTAAGAGCTGGCTCAATGGTTCAGGC 838
DB 664 TCTCAGCCCTGTACTGCTTCAACTATGAGTGTGTAAGAGCTGGCTCAATGGTTCAGGC 605
QY 839 CGAAGACCAAGACTTCTGTGGGATGAGCTTGTGGCTGGTGGATCTCAGGACGGTGG 898
DB 604 CGAAGACCAAGACTTCTGTGGGATGAGCTTGTGGCTGGTGGATCTCAGGACGGTGG 545
QY 899 CTCAGTGTGACTTACCTTTGACGTGTGTAAGACCAAGCCAGCTGCTCTGGGAG 958
DB 544 CTCAGTGTGACTTACCTTTGACGTGTGTAAGACCAAGCCAGCTGCTCTGGGAG 485
QY 959 CGATGAGGCTGTGAGTGAACCCCTTGCATGTGACTTCCACCTGGTGTGCTGCGGA 1018
DB 484 CGATGAGGCTGTGAGTGAACCCCTTGCATGTGACTTCCACCTGGTGTGCTGCGGA 425
QY 1019 GGATCGGGCGAGTGGGCGACCAAGGAGCTCTTTTCAGGCTTCTTCTCGATCATCA 1078
DB 424 GGATCGGGCGAGTGGGCGACCAAGGAGCTCTTTTCAGGCTTCTTCTCGATCATCA 365
QY 1079 AGGTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1138
DB 364 AGGTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
QY 1139 AGAGGCTGAACCAAGGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGGAGCAAGCCCG 1198
DB 304 AGAGGCTGAACCAAGGACCGGCTTCTGGGCGGCTGAAAGGGGCAAGGAGCAAGCCCG 245

QY 1199 TCTCTCCACGATGGGAGAGGGCAGGAGAGACCAGCAAGTGCCTTTTCTCAGCA 1258
DB 244 TTTCTCTCCACGATGGGAGAGGGCAGGAGAGACCAGCAAGTGCCTTTTCTCAGCA 185
QY 1259 CTGAGGAGGGGGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1318
DB 184 CTGAGGAGGGGGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 125
QY 1319 CTCTGGGCGGCCAGCAGCTTCTCTCAGACAGCTTCTCTCTCTCTCTCTCTCTCTCT 1378
DB 124 TTTCTGGGCGGCCAGCAGCTTCTCTCAGACAGCTTCTCTCTCTCTCTCTCTCTCTCT 65
QY 1379 TCATCACTTACCCACCCCAAGTTCAGACCAAAATCTTCCAGCTGCCCCCTTCTGTTT 1438
DB 64 TCATCACTTACCCACCCCAAGTTCAGACCAAAATCTTCCAGCTGCCCCCTTCTGTTT 5
QY 1439 CCCT 1442
DB 4 CCCT 1

RESULT 12

BM915493
LOCUS BM915493
DEFINITION AGENCOURT_6701631 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5481679
5', mRNA sequence.

ACCESSION BM915493
VERSION BM915493.1 GI:19365872

KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 959)

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2006 row: i column: 08
High quality sequence stop: 687.

FEATURES

source

1. .959
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5481679"
/clone.lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 162 a 308 c 290 g 199 t

ORIGIN

Query Match 55.4%; Score 855.8; DB 14; Length 959;

Best Local Similarity 96.7%; Pred. No. 1.5e-199;

Matches 916; Conservative 0; Mismatches 27; Indels 4; Gaps 4;

QY 14 GCGAGCGCGGAGCCAGGAGGGCGGAGCGGACGTAATGGCCCGCTTGGCCCTGGGGCC 73

```

Db 4 GGGCGCGCGGACCGCGCACAGGCGGCGGAGCCGGTATGGGCGCCCGCTGACCCCTGGCGGCC 63
QY 74 GGGCGCGCGGACCGCGCACAGGCGGAGCCGGTATGGGCGCCCGCTGACCCCTGGCGGCC 133
Db 64 GGGCGCGCGGACCGCGCACAGGCGGAGCCGGTATGGGCGCCCGCTGACCCCTGGCGGCC 123
QY 134 CTGGGGGATCAGCGCCCTCCAGCAATGCTGGGCTCAGGACCGGGGCTGGTGTACT 193
Db 124 CTGGGGGATCAGCGCCCTCCAGCAATGCTGGGCTCAGGACCGGGGCTGGTGTACT 183
QY 194 CTCTCTTATGACACCCCTGAGCGTGGTGAAGGTTTCGGCTGCAGTCTCAGCGCCCTCA 253
Db 184 CTCTCTTATGACACCCCTGAGCGTGGTGAAGGTTTCGGCTGCAGTCTCAGCGCCCTCA 243
QY 254 TGCGCAGCGAGCTGATGCTCTCTCAGACGCTGGAGGCTCTCTATACCAATGGAAGT 313
Db 244 TGCGCAGCGAGCTGATGCTCTCTCAGACGCTGGAGGCTCTCTATACCAATGGAAGT 303
QY 314 GCTCTCTGTAATGCTGCTCTGGAGCTCTCTGAGGCTCTCTGAGGCTCTCTGAGGCT 373
Db 304 GCTCTCTGTAATGCTGCTCTGGAGCTCTCTGAGGCTCTCTGAGGCTCTCTGAGGCT 363
QY 374 GTGCCACCTGTTTCAAGACCCCTACCGCTTCACTGGCACCATGGATGCTTCTGTAAGA 433
Db 364 GTGCCACCTGTTTCAAGACCCCTACCGCTTCACTGGCACCATGGATGCTTCTGTAAGA 423
QY 434 TCGTGAGGACAGGCGCACAGGACCCCTCTGGAGCGGCTCTCCCGCACCCCTGGTGATGA 493
Db 424 TCGTGAGGACAGGCGCACAGGACCCCTCTGGAGCGGCTCTCCCGCACCCCTGGTGATGA 483
QY 494 CTGTGCGAGCTACCGCACTACTTCACTGCTATGACCACTGAAGGCTTCTGTTGTG 553
Db 484 CTGTGCGAGCTACCGCACTACTTCACTGCTATGACCACTGAAGGCTTCTGTTGTG 543
QY 554 GTGAGGCGCTGACCTCTGACCTCTACGACCCATGGTGGCTGGCGGCTGGCCCGCTGG 613
Db 544 GTGAGGCGCTGACCTCTGACCTCTACGACCCATGGTGGCTGGCGGCTGGCCCGCTGG 603
QY 614 GCACCGTGACTGTGATCAGCCCGCTGGAGCTTATCGCGACAAAGCTCGAGGCTCAGCATG 673
Db 604 GCACCGTGACTGTGATCAGCCCGCTGGAGCTTATCGCGACAAAGCTCGAGGCTCAGCATG 663
QY 674 TGCTGTACCGGAGCTGGGTGCTGTGTGGAATGTCAGTGGCTCAGGCTGGCTGGGCT 733
Db 664 TGCTGTACCGGAGCTGGGTGCTGTGTGGAATGTCAGTGGCTCAGGCTGGCTGGGCT 723
QY 734 CACTGTGGCTGGGCTGGGCGCCCTGCTGCGATGTCGAGATGTCCTTCTCAGGCTGTACT 793
Db 724 CACTGTGGCTGGGCTGGGCGCCCTGCTGCGATGTCGAGATGTCCTTCTCAGGCTGTACT 783
QY 794 GGTTCACCTATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTCAGGCGAAGGACGAGCTT 853
Db 784 GGTTCACCTATGAGCTGGTGAAGAGCTGGCTCAATGGGTTTCAGGCGAAGGACGAGCTT 843
QY 854 CTGTGGGATGAGCTTGTGGCTGGTGCA-TCTCAGGAGCTT-GGCTGAGTGTGA 910
Db 844 CTGTGGGATGAGCTTGTGGCTGGTGCA-TCTCAGGAGCTTTCAGGAGCTTGGCTGCTGA 903
QY 911 CTCTACCC-TTTGAGTGGTAAAGACCCCAAGCCAGGCTGCTCTGGG 956
Db 904 CTCTACCC-TTTGAGTGGTAAAGACCCCAAGCCAGGCTGCTCTGGG 950

RESULT 13
BQ949458 939 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8824335 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6459569
5', mRNA sequence.
ACCESSION BQ949458
VERSION BQ949458.1 GI:22364936
KEYWORDS EST.
SOURCE human.

```

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 939)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DP/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILLNL at:
<http://image.llnl.gov>
Plate: LICM2644 row: n column: 18
High quality sequence stop: 687.

FEATURES

source

1..939
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6459569"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
ECORI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by King Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 154 a 300 c 294 g 191 t

Query Match

Best Local Similarity 55.2%; Score 854; DB 14; Length 939;

Matches 905; Conservative 0; Mismatches 25; Indels 6; Gaps 3;

```

QY 1 GCTAGGTGCTCGAGCGCGG----GAGCCACGAGGCGGCGAGCTAATGGGCC 56
Db 4 GCCTAGGTGCTCGAGCGCGCGGACCGGCGACAGCGCGGCGGCGGATGGGCC 63
QY 57 GCCTGGCCCTGGGCGCGCGCCCGACGACGACGCTAGAGCAGACTGAAGCTCAA 116
Db 64 GCCTGACCTGGGCGCGCGCGCGACGACGACGCTAGAGCAGACTGAAGCTCAA 123
QY 117 GATGCTGACGAGGACCTCGGGGATCAGCCCTCCAGCAATGTCGCTCAGGCAC 176
Db 124 GATGCTGACGAGGACCTCGGGGATCAGCCCTCCAGCAATGTCGCTCAGGCAC 183
QY 177 CGGGCTGTGTACTCTCTCTTCATGACACCCCTGGAGCTGGTGAAGCTTCGCTGCA 236
Db 184 CGGGCTGTGTACTCTCTCTTCATGACACCCCTGGAGCTGGTGAAGCTTCGCTGCA 243
QY 237 GTCCTCAGGGGCTTCATGCGGACGAGCTAGTCTCTCCAGACTGTGGAGCTCTC 296
Db 244 GTCCTCAGGGGCTTCATGCGGACGAGCTAGTCTCTCCAGACTGTGGAGCTCTC 303
QY 297 CTATACCAATGGAAGTGGCTCTGTATTGCAATGGTGTCTCGAGCCCTCTGTACTGTG 356
Db 304 CTATACCAATGGAAGTGGCTCTGTATTGCAATGGTGTCTCGAGCCCTCTGTACTGTG 363
QY 357 CCCAAATGGTGCCCGCTGTGCCACCTGTGTTTCAAGACCCCTACCCGCTTCACCTGGACCAT 416
Db 364 CCCAAATGGTGCCCGCTGTGCCACCTGTGTTTCAAGACCCCTACCCGCTTCACCTGGACCAT 423
QY 417 GGATGCTTGTGAAGATGTCGAGGACGAGGACCCAGGACCTCTGTGAGGCGCTTCCC 476
Db 424 GGATGCTTGTGAAGATGTCGAGGACGAGGACCCAGGACCTCTGTGAGGCGCTTCCC 483

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 22, 2003, 19:04:17 ; Search time 3022.46 Seconds
(without alignments)
10726.543 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagcttcaagatggctg.....aggggcaaggaggaaggac 1114

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
To: number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1114	100.0	1114	6	AX403084 Sequence
2	1114	100.0	1114	9	AF317711 Homo sapi
3	1056	94.8	1550	9	AF317711 Homo sapi
4	1056	94.8	1550	9	AF317711 Homo sapi
5	1054.4	94.6	1546	6	AX403085 Sequence
6	1054.4	94.6	1556	9	AF151827 Homo sapi
7	1054.4	94.6	1579	9	BC001398 Homo sapi
8	1054.4	94.6	1643	6	AF116038 Sequence
9	1052.8	94.5	1554	9	AK026060 Homo sapi
10	1002.4	90.0	1402	17	AF119864 AF119864 Homo sapi
11	651	58.4	173434	2	AC098934 Homo sapi
12	560.8	50.3	128802	2	AC123139 Rattus no
13	558.8	50.2	55252	2	AC1194893 Continuation (4 of
14	554	49.7	156175	2	AC111356 Rattus no
15	552.8	49.6	195701	2	AC124203 Rattus no
16	488.8	43.9	177613	2	AC123014 Rattus no
17	472	42.4	704	6	AX305924 Sequence
18	448.4	40.3	758	6	AX421524 Sequence
19	440	39.5	194025	2	AC114721 Rattus no
20	298.4	26.8	138839	2	AC108295 Rattus no
21	242.2	21.7	1284	9	BC027322 Homo sapi
22	240.6	21.6	1517	9	AF125531 Homo sapi
23	210.4	18.9	162456	2	AC099097 Rattus no
24	176.4	15.8	139488	9	AC003043 Homo sapi
25	176.4	15.8	180520	2	AC019152 Homo sapi
26	175.2	15.7	195768	2	AC068775 Homo sapi
27	174.8	15.7	55774	2	AC091303 Homo sapi
28	170.2	15.3	512	6	AX400653 Sequence
29	170	15.3	537	6	AX106606 Sequence
30	170	15.3	537	6	AX140897 Sequence
31	170	15.3	537	6	AX200757 Sequence
32	170	15.3	537	6	AX267413 Sequence
33	168.2	15.1	82586	2	AC096040 Rattus no
34	168.2	15.1	188670	2	AC127606 Rattus no
35	156	14.0	194025	2	AC114721 Rattus no
36	142	12.7	18391	2	AC019151 Homo sapi
37	140.4	12.6	165490	2	AC025326 Homo sapi
38	140.4	12.6	190430	2	AC087873 Mus muscu
39	140.4	12.6	219037	10	AL596258 Mouse DNA
40	139.6	12.5	197197	2	AC118439 Rattus no
41	136.4	12.2	217422	2	AC116713 Mus muscu
42	131	11.8	3070	3	AY070516 Drosophil
43	130.4	11.7	1463	3	AY051544 Drosophil
44	130.4	11.7	16791	2	AC018172 Drosophil
45	130.4	11.7	175041	3	AC008092 Drosophil

ALIGNMENTS

RESULT 1
AX403084
LOCUS AX403084 1114 bp DNA linear PAT 07-JUN-2002
DEFINITION Sequence 1 from Patent WO0198355.
ACCESSION AX403084
VERSION AX403084.1 GI:21388033
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Lewin, D., Adams, S.H. and Yu, X.X.
TITLE Cgi-69 compositions and methods of use
JOURNAL Patent: WO 0198355-A 1 27-DEC-2001;

Genentech, Inc. (US) : Curagen Corporation (US)

FEATURES
source
1. .1114
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 194 a 349 c 332 g 239 t
ORIGIN

Query Match 100.0%; Score 1114; DB 6; Length 1114;
Best Local Similarity 100.0%; Pred. No. 5.4e-244;
Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGCTTCAAGTGGCTGACGAGACCTGGGGCATCAGCCCTCCAGAAATGGT 60
DB 1 CTGAAGCTTCAAGTGGCTGACGAGACCTGGGGCATCAGCCCTCCAGAAATGGT 60
QY 61 GGCTCAGGACCGGGGGTGGTGTACCTCTCTCTCATGACACCCCTGGACGTGTGAA 120
DB 61 GGCTCAGGACCGGGGGTGGTGTACCTCTCTCTCATGACACCCCTGGAGTGTGAA 120
QY 121 GGTTCGGCTGCACTCTCAGCGGCTCTCATGGCAGGAGCTGATGCTCTCTCCAGACT 180
DB 121 GGTTCGGCTGCACTCTCAGCGGCTCTCATGGCAGGAGCTGATGCTCTCTCCAGACT 180
QY 181 GTGAGGCTCTCTTATACCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
DB 181 GTGAGGCTCTCTTATACCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 241 GTATTGCAATGGTGTCTCTGAGGCTCTCTACCTGTGCCCCAATGTGCGCGCTGCCCAC 300
DB 241 GTATTGCAATGGTGTCTCTGAGGCTCTCTACCTGTGCCCCAATGTGCGCGCTGCCCAC 300
QY 301 CTGGTTTCAAGACCTTACCGCTTCACTGCGACCATGATGCTCTCTGAGAGTGTGAG 360
DB 301 CTGGTTTCAAGACCTTACCGCTTCACTGCGACCATGATGCTCTCTGAGAGTGTGAG 360
QY 361 GCACGGGACACGAGGACCTCTGAGGCGCTCTCCCGCCACCTGGTGTGATGCTGTC 420
DB 361 GCACGGGACACGAGGACCTCTGAGGCGCTCTCCCGCCACCTGGTGTGATGCTGTC 420
QY 421 AGCTACCGGCTTACTTCACTGCTTATGACCAATGAGGCTTCTCTGTGTGCTGAGC 480
DB 421 AGCTACCGGCTTACTTCACTGCTTATGACCAATGAGGCTTCTCTGTGTGCTGAGC 480
QY 481 CTTGACCTTGAATCTTACGACCATGCTGCTGGCGCTGCGCGCTGGCACCGT 540
DB 481 CTTGACCTTGAATCTTACGACCATGCTGCTGGCGCTGCGCGCTGGCACCGT 540
QY 541 GACTGTGATCAGGCTTATGAGGCTTATGAGCAAAAGCTGAGGCTCAGCATGTGTGTA 600
DB 541 GACTGTGATCAGGCTTATGAGGCTTATGAGCAAAAGCTGAGGCTCAGCATGTGTGTA 600
QY 601 CCGGAGTGGGTGCTGTGTTCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CCGGAGTGGGTGCTGTGTTCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 GTGGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GTGGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 CTATGAGTGTGAGAGCTGGGTCTAATGGCTCAGGCGGAGGACGAGCTTCTGTGG 780
DB 721 CTATGAGTGTGAGAGCTGGGTCTAATGGCTCAGGCGGAGGACGAGCTTCTGTGG 780
QY 781 CATGAGCTTTGGCTGGTGGTATCTCAGGAGCGGTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 CATGAGCTTTGGCTGGTGGTATCTCAGGAGCGGTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 TGACGTGTGTAAGACCCAAAGCTGCTCTGAGGAGGATGAGGCTGTGAGAGTGA 900
DB 841 TGACGTGTGTAAGACCCAAAGCTGCTCTGAGGAGGATGAGGCTGTGAGAGTGA 900
QY 901 CCCCCTGCTATGTGACTCCACCTGGTGTGCTGCGAGGATCCGGGCGGAGTCCGGCAC 960

DB 901 CCCCCTGCTATGTGACTCCACCTGGTGTGCTGCGAGGATCCGGCGGAGTCCGGCAC 960
QY 961 CAAGGACTCTTTGAGGCTTCTCTCTCGGATCATCAAGGCTGCGCTCTCTGTCAT 1020
DB 961 CAAGGACTCTTTGAGGCTTCTCTCTCGGATCATCAAGGCTGCGCTCTCTGTCAT 1020
QY 1021 CATGATCAGCACCCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGCTGAACACCGGCT 1080
DB 1021 CATGATCAGCACCCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGCTGAACACCGGCT 1080
QY 1081 TCTGGCGGTGAAGGGGGAAGGAGGAGGAC 1114
DB 1081 TCTGGCGGTGAAGGAGGAGGAGGAC 1114
RESULT 2
AF317711 1114 bp mRNA linear PRI 13-DEC-2001
LOCUS Homo sapiens mitochondrial carrier protein CGI-69 long form mRNA,
DEFINITION complete cds.
ACCESSION AF317711
VERSION AF317711.1 GI:12620399
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1114)
AUTHORS Yu, X.X., Lewin, D.A., Zhong, A., Brush, J., Schow, P.W., Sherwood, S.W.,
Pan, G., and Adams, S.H.
TITLE Overexpression of the human 2-oxoglutarate carrier lowers
mitochondrial membrane potential in HEK-293 cells: contrast with
the unique cold-induced mitochondrial carrier CGI-69
JOURNAL Biochem J. 353 (Pt 2), 369-375 (2001)
MEDLINE 21066063
PUBMED 11139402
REFERENCE 2 (bases 1 to 1114)
AUTHORS Yu, X.X., Lewin, D.A., Zhong, A., Brush, J., Schow, P.W., Sherwood, S.W.,
Pan, G., and Adams, S.H.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2000) Endocrinology, Genentech, Inc., 1 DNA Way,
South San Francisco, CA 94080, USA
FEATURES
source
1. .1114
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue type="liver"
14. .1093
/note="similar to CGI-69 encoded by GenBank Accession
Number AF151827"
/product="mitochondrial carrier protein CGI-69 long form"
/protein_id="AAG60687.1"
/db_xref="GI:12620400"
/translation="MADQDPAGISPLQWVASTGAVTSLFWTLDVVKVLSORP
SMASELSPSSRLWSLSYTKPSQSTKCLLYCNGLVLEPLVLPNGACATWPOPT
RFTGTDAFVKVIRHEGTRILWSGLPATLVNTVPATLYFTADQLKAFCLGRALISD
LYAPVAGALARIQTVTVISPLELMRTKLQAHVSYRELGCACVRTAVAGQWRLMLG
WGPTALRDVPFSALVWFNLYVKSNLGRPKDQTSVGMFSVAGGISGTVAAVLTLPF
DIVKTORVALGMEAVRNVLHVDSTWLLLRIRAESGTLKGLFGLPLRIKAAPSC
ALMISTYFKSFFGLNQDRLLGG"

BASE COUNT 194 a 349 c 332 g 239 t
ORIGIN
Query Match 100.0%; Score 1114; DB 9; Length 1114;
Best Local Similarity 100.0%; Pred. No. 5.4e-244;
Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGAGCTTCAAGTGGCTGACGAGACCTGGGCTATCAGCCCTCCAGAAATGGT 60
DB 1 CTGAGCTTCAAGTGGCTGACGAGACCTGGGCTATCAGCCCTCCAGAAATGGT 60

QY 61 GGCCTCAGGACCGGGCTGTGGTACCTCTCTCTCATGACACCCCTGGAGTGGTAA 120
Db 61 GGCCTCAGGACCGGGCTGTGGTACCTCTCTCTCATGACACCCCTGGAGTGGTAA 120
QY 121 GGTTCGCTCGAGTCTCAGCGCCCTCATGCGCAGGAGCTATGCTTCTCCAGACT 180
Db 121 GGTTCGCTCGAGTCTCAGCGCCCTCATGCGCAGGAGCTATGCTTCTCCAGACT 180
QY 181 GTGAGGCTCTCTATACAAATTCCTCTCTCCATCCACAGGGAAGTGCCTCT 240
Db 181 GTGAGGCTCTCTATACAAATTCCTCTCTCCATCCACAGGGAAGTGCCTCT 240
QY 241 GTATTGAATGTGTCTCGAGCTCTGTACTGTGCCAAATGCTGCCGCTGGCCAC 300
Db 241 GTATTGAATGTGTCTCGAGCTCTGTACTGTGCCAAATGCTGCCGCTGGCCAC 300
QY 301 CTGGTTTCAGACCTACCGCTTCACTGCGACCATGATGCTTGTGAAGATCGTGAG 360
Db 301 CTGGTTTCAGACCTACCGCTTCACTGCGACCATGATGCTTGTGAAGATCGTGAG 360
QY 361 GCACGAGGACAGGACCTCTGGAGCGGCTCTCCCGCCACCTGGTGATGACTGTGCC 420
Db 361 GCACGAGGACAGGACCTCTGGAGCGGCTCTCCCGCCACCTGGTGATGACTGTGCC 420
QY 421 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGGTGGC 480
Db 421 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGGTGGC 480
QY 481 CTTGACCTCTACCTGACGACCATGCTGGCTGGCGGCTGGCGGCTGGCGGCTGGC 540
Db 481 CTTGACCTCTACCTGACGACCATGCTGGCTGGCGGCTGGCGGCTGGCGGCTGGC 540
QY 541 GACTGTGATCAGCCCTCTGAGGCTTATCGGACAAAGCTCGAGGCTCAGCATGTGTGTA 600
Db 541 GACTGTGATCAGCCCTCTGAGGCTTATCGGACAAAGCTCGAGGCTCAGCATGTGTGTA 600
QY 601 CCGGAGCTGGTGGCTGTGGTTCGAACTGCAAGTGGCTCAGGCTGGCTCAGTGTG 660
Db 601 CCGGAGCTGGTGGCTGTGGTTCGAACTGCAAGTGGCTCAGGCTGGCTCAGTGTG 660
QY 661 CTGGGCTGGGCGCCACCTGCGCTCGAGATGTCCTTCTCAGCCCTGACTGTTCAA 720
Db 661 CTGGGCTGGGCGCCACCTGCGCTCGAGATGTCCTTCTCAGCCCTGACTGTTCAA 720
QY 721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGG 780
Db 721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGG 780
QY 781 CATGAGCTTGTGGCTGGTGGCTCTCAGGAGCGTGGCTGAGTGTGCTTACCTT 840
Db 781 CATGAGCTTGTGGCTGGTGGCTCTCAGGAGCGTGGCTGAGTGTGCTTACCTT 840
QY 841 TGAGCTGGTAAAGACCAAGCGAGTGGCTCTGGGAGCGATGGAGGCTGAGAGTAA 900
Db 841 TGAGCTGGTAAAGACCAAGCGAGTGGCTCTGGGAGCGATGGAGGCTGAGAGTAA 900
QY 901 CCCCTTCATGTGGATCCACTGGCTCTGCTCGGAGGATCCGGCGGAGTCCGGCAC 960
Db 901 CCCCTTCATGTGGATCCACTGGCTCTGCTCGGAGGATCCGGCGGAGTCCGGCAC 960
QY 961 CAAGGACTCTTGGAGGCTTCTTCTCGGATCATCAGGCTGCCCCCTCTGTGGCAT 1020
Db 961 CAAGGACTCTTGGAGGCTTCTTCTCGGATCATCAGGCTGCCCCCTCTGTGGCAT 1020
QY 1021 CATGATCAGCACCTATGACTTCGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCT 1080
Db 1021 CATGATCAGCACCTATGACTTCGCAAAAGCTTCTTCAGAGGCTGAACAGGACCGGCT 1080
QY 1081 TCTGGGCGCTGAAGGGGCAAGGAGGCAAGGAC 1114
Db 1081 TCTGGGCGCTGAAGGGGCAAGGAGGCAAGGAC 1114

RESULT 3
HSM801436 1550 bp mRNA linear PRI 10-MAR-2001
LOCUS Homo sapiens mRNA; cDNA DKFZp434C119 (from clone DKFZp434C119);
DEFINITION complete cds.
ACCESSION AL133584
VERSION AL133584.1 GI:6599169
KEYWORDS human
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1550)
AUTHORS Wiemann,S., Well,B., Wellenreuther,R., Gassenhuber,J., Glassl,S.,
Ansorci,W., Boecher,M., Bloeker,H., Bausechs,S., Blum,H.,
Lauber,J., Duesterhoeft,A., Seyer,A., Koehrer,K., Strack,N.,
Mewes,H.W., Othenwaelder,B., Obermaier,B., Tampe,J., Heubner,D.,
Wambutt,R., Korn,B., Klein,M. and Poustka,A.
Toward a Catalog of Human Genes and Proteins: Sequencing and
Analysis of 500 Novel Complete Protein Coding Human cDNAs
Genome Res. 11 (3), 422-435 (2001)
JOURNAL 11230166
PUBMED 2 (bases 1 to 1550)
REFERENCE Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
AUTHORS Direct Submission
TITLE Submitted (15-DEC-1999) MIPS, Am Klopferspitz 18a, D-82152
JOURNAL Martinried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp434C119) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..1550
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="DKFZp434C119"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSPORTi; host
DH10B; sites: NotI + SalI"
/dev_stage="adult"
118..1173
/gene="DKFZp434C119"
118..1173
/gene="DKFZp434C119"
/note="CGI-69 protein, strong similarity to mitochondrial
carrier proteins"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB63728.1"
/db_xref="GI:6599170"
/db_xref="SPTREMBL:O9UF66"
/translation="MADQDPAGISPLQKQVAGSGAVVTSLFMTPLDVKVRLQSQRP
SKVSEMFSSRLWSLSTYKWKLLYCNGVLEPLCPNGARCATFQDPTFTGTMDA
FYKIVRHGRTLMGSLPATLVMTATYATYDOLKAPICGRALTSILYAPMVG
ALARGTIVTISPLELMRTKLOAHVSVRELGCACVTAQAGGWSRLWGPTALRD
VPFSALYVFNVELVKSLNGLRPKQTSVGMSEVAGISGTVAAILTLFDFVVKTRQ
VALGMEAVRVNPLHVDSTWLLLRIRAESGKGLFAGFPRITKAAPSCAIMISTYE
FKSFFQLNQDRLGG"
polyA_site 1540
BASE COUNT 278 a 490 c 459 g 323 t
ORIGIN
Query Match 94.8%; Score 1056; DB 9; Length 1550;
Best Local Similarity 97.8%; Pred. No. 9.3e-231;
Matches 1090; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 CTGAGGCTTCAAGATGCTGATCCAGGACCTTGGGATCAGCCCTCCAGCAATGGT 60
 Db 105 CTGAAGCTTCAAGATGCTGATCCAGGACCTTGGGATCAGCCCTCCAGCAATGGT 164
 QY 61 GGCCTCAGGACCGGGGCTGTGGTACTCTCTCTCATCACACCCCTGGACGTGTGAA 120
 Db 165 GGCCTCAGGACCGGGGCTGTGGTACTCTCTCTCATCACACCCCTGGACGTGTGAA 224
 QY 121 GGTTCGCTGAGTCTCAGCGGCTTCATGGCCAGGAGCTGATGCTCTCTCCAGACT 180
 Db 225 GGTTCGCTGAGTCTCAGCGGCTTCATGGCCAGGAGCTGATGCTCTCTCCAGACT 284
 QY 181 GTGAGGCTCTCTATACCAATGGCTCTCTCTCTCAATCCAGGAAAGTCCCTCT 240
 Db 285 GTGAGGCTCTCTATACCAAT-----GGAAGTCCCTCTCT 320
 QY 241 GTATGCAATGGTCTGAGGCTCTGTACTGTGCCAAATGTGCGCCCTCTGCGAC 300
 Db 321 GTATGCAATGGTCTGAGGCTCTGTACTGTGCCAAATGTGCGCCCTCTGCGAC 380
 QY 301 CTGGTTTCAAGACCTTACCGCTTCACTGGCAGCATGGATGCTCTGAGAGATCGTGAG 360
 Db 381 CTGGTTTCAAGACCTTACCGCTTCACTGGCAGCATGGATGCTCTGAGAGATCGTGAG 440
 QY 361 GCACGAGGACACGAGGACCTCTGAGCGGCTCCCGGACCCCTGGTGTGATGCTGTGCC 420
 Db 441 GCACGAGGACACGAGGACCTCTGAGCGGCTCCCGGACCCCTGGTGTGATGCTGTGCC 500
 QY 421 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTCTCTGTGTGTGAGC 480
 Db 501 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTCTCTGTGTGTGAGC 560
 QY 481 CTTGACCTCTGACCTTACGACCATGGTGGCTGGCGGCTGGCGGCTGGCGGACCGT 540
 Db 561 CTTGACCTCTGACCTTACGACCATGGTGGCTGGCGGCTGGCGGCTGGCGGACCGT 620
 QY 541 GACTGTGATCAGCGGCTTATCGGACAAAGCTTCAGGCTTCAGGATGTGTGCTGA 600
 Db 621 GACTGTGATCAGCGGCTTATCGGACAAAGCTTCAGGCTTCAGGATGTGTGCTGA 680
 QY 601 CCGGAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 681 CCGGAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 740
 QY 661 GCTGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db 741 GCTGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
 QY 721 CTATGAGCTGGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Db 801 CTATGAGCTGGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860
 QY 781 CATGAGCTTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db 861 CATGAGCTTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920
 QY 841 TCAGTGTGAAGACCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Db 921 TGAGTGTGAAGACCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980
 QY 901 CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db 981 CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1040
 QY 961 CAAGGAGCTTTGTGAGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 1041 CAAGGAGCTTTGTGAGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
 QY 1021 CATGATCAGGACCTTATGAGTTCGGAAGGCTTCTTCCAGAGCTGCAACGAGCCGCT 1080
 Db 1101 CATGATCAGGACCTTATGAGTTCGGAAGGCTTCTTCCAGAGCTGCAACGAGCCGCT 1160
 QY 1081 TCTGGGCTGCTGAAAGGGGCAAGGAGGCAAGGAC 1114

Db 1161 TCTGGGCTGCTGAAAGGGGCAAGGAGGCAAGGAC 1194

RESULT 4
 BC009330
 LOCUS 1560 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Homo sapiens, clone MGC:16684 IMAGE:4126268, mRNA, complete cds.
 ACCESSION BC009330
 VERSION BC009330.1 GI:14424605
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1560)
 Direct Submission
 Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc.mgc@nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 26 Row: g Column: 12.
 FEATURES
 source 1..1560
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:16684 IMAGE:4126268"
 /issue_type="Brain, neuroblastoma"
 /clone_lib="NIH MGC 19"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 103..1158
 /codon_start=1
 /product="Unknown (protein for MGC:16684)"
 /protein_id="AAH09330.1"
 /db_xref="GI:14424606"
 /translation="MADDPAGISPLQOMVASGTGAVVTSLEMTPLDVVKVRLQSORP SVASELMSSRLWSLSTYTKKKLLCYNGVLEPLYLCPNGARCATWFOPTRTGTMDA FVKVVRHGSTRILWSGFPATLVNVPATAYFYFDQKAFGLCGRALTSDLVAPWVAG ALARLGTVTIVSPLNMRKTLQHVSVRELACVATVAGQGWRSILWLGWGTALRD VPFSLYVFNVELVKSLNGLRPKDQTSVSGMSFVAGISGTVAALVTLPPDVVKTORQ VALGAEAVRVNPLHVDSTWLLLRIRABSGTKGLFAGLFPRIIAKAPSCAIIISTYE FCKSPFORLNQDRLLGG"
 BASE COUNT 296 a 488 c 453 g 323 t
 ORIGIN
 Query Match 94.8%; Score 1056; DB 9; Length 1560;
 Best Local Similarity 97.8%; Pred. No. 9.3e-231;
 Matches 1090; Conservative 0; Mismatches 0; Indels 24; Gaps 1;
 QY 1 CTGAAGCTTCAAGATGCTGATCCAGGACCTTGGGATCAGCCCTCCAGCAATGGT 60

[illegible]

Db 1146 TCTGGCGCGCTGAAGAGGCAAGGAGCAAGGAC 1179

RESULT 5
AX403085

LOCUS AX403085 1546 bp DNA linear PAT 07 JUN 2002

DEFINITION Sequence 2 from Patent WO0198355.

ACCESSION AX403085

VERSION AX403085.1

KEYWORDS Gi:21388034

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Lewin, D., Adams, S.H. and Yu, X.X.
TITLE Cgi-69 compositions and methods of use
JOURNAL Genet. WO 0198355-A 27-DEC-2001;
Genentech, Inc. (US); Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..1546
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 272 a 489 c 459 g 326 t

ORIGIN

Query Match 94.6%; Score 1054.4; DB 6; Length 1546;
Best Local Similarity 97.8%; Pred. No. 2.2e-230;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

Qy 1 CTGAAGCTTCAAGATGGCTGACCAGGACCCCTGCGGGCATCAGCCCCCTCCAGCAAAATGCT 60

Db 105 CTGAAGCTTCAAGATGGCTGACCAGGACCCCTGCGGGCATCAGCCCCCTCCAGCAAAATGCT 164

Qy 61 GGCCTCAGGACACGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGAGCACTGTGAA 120

Db 165 GGCCTCAGGACACGGGGCTGTGGTTACTCTCTCTTCATGACACCCCTGAGCACTGTGAA 224

Qy 121 GGTTCGCTGCAGTCTCAGCGCCCTCCATGGCCAGGAGCTGATGCTTCTCCAGACT 180

Db 225 GGTTCGCTGCAGTCTCAGCGCCCTCCATGGCCAGGAGCTGATGCTTCTCCAGACT 284

Qy 181 GTGAGGACCTCTCCTATACCAAAATGCCCCTCTCTCTCCAATCCACGAGGAATGCCCTCT 240

Db 285 GTGAGGACCTCTCCTATACCAAT-----GGAGTGGCTCTCT 320

Qy 241 GTATTGMAATGGTGTCTGGAGCCCTGTACTGTGCCCAAAATGTCGCCCTGTGCCAC 300

Db 321 GTATTGCAATGGTGTCTGGAGCCCTGTACTGTGCCCAAAATGTCGCCCTGTGCCAC 380

Qy 301 CTGTTTTCAGAGCCCTACCGCTTCCACTGGCACCATTGGATGCTTCGTGAAGATCGTAG 360

Db 381 CTGGTTTTCAGAGCCCTACCGCTTCCACTGGGACCAATGGATGCTTCGTGAAGATCGTAG 440

Qy 361 GCACGAGGCGCACGAGGACCCCTCTGGAGCGGCTCCCGCCACCCCTGGTGATGACTGTGCC 420

Db 441 GCACGAGGCGCACGAGGACCCCTCTGGAGCGGCTCCCGCCACCCCTGGTGATGACTGTGCC 500

Qy 421 AGCTACCGCACTATCTTCATCTGCCATTATGACCAACTGAAGGCTTCTCTGTGGTCGAGC 480

Db 501 AGCTACCGCACTATCTTCATCTGCCATTATGACCAACTGAAGGCTTCTCTGTGGTCGAGC 560

Qy 481 CTTGACCTCTGACCTCTACGACACCCATTGGTGGCGGCTGGCCGCTGGGACACCGT 540

Db 561 CTTGACCTCTGACCTCTACGACACCCATTGGTGGCGGCTGGCCGCTGGGACACCGT 620

Qy 541 GACTGTGATCAGCGCCCTCGAGCTTATGCGGCAAAAGCTGCAAGGCTCAGCATGTGCTGTA 600

Db 621 GACTGTGATCAGCGCCCTCGAGCTTATGCGGCAAAAGCTCAGAGCTCAGCATGTGCTGTA 680

Qy 601 CCGGAGCTGGGTGCTCTGTGTTTCGAATGTCAGTGGCTCAGGCTGGCTGCGCTCACTGTG 660

Db 681 CCGGAGCTGGGTGCTCTGTGTTTCGAATGTCAGTGGCTCAGGCTGGCTGCGCTCACTGTG 740

Db 1041 CAAGGAGCTCTTTCAGGCTTCTTCTCGATCATCAAGGCTGCCCCCTCTCTGTCGCT 1100
 QY 1021 CATGATCAGCACCTATGAGTTGGGAAAAGCTTCTTCCAGAGGCTGAACCCAGGACCGGCT 1080
 Db 1101 CATGATCAGCACCTATGAGTTGGGAAAAGCTTCTTCCAGAGGCTGAACCCAGGACCGGCT 1160
 QY 1081 TCTGGCGGCTGAAGGGGCAAGGCAAGGAC 1114
 Db 1161 TCTGGCGGCTGAAGGGGCAAGGCAAGGAC 1194

RESULT 7
 BC001398 1579 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, CGI-69 protein, clone MGC:1163 IMAGE:3051436, mRNA,
 DEFINITION complete cds.
 ACCESSION BC001398
 VERSION BC001398.1 GI:12655092
 KEYWORDS MGC.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1579)
 Direct Submission
 Submitted (12-DEC-2000) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
 Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 4 Row: o Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10438782.
 Location/Qualifiers
 1. .1579
 /organism="Homo sapiens"
 /db_xref="locusID:51629"
 /db_xref="taxon:9606"
 /clone="MGC:1163 IMAGE:3051436"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NIH_MGC_15"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 131. 1186
 /codon_start=1
 /product="CGI-69 protein"
 /protein_id="AAH01398.1"
 /db_xref="GI:12655093"
 /translation="MADQDPAIGSLPQOMVAGSTGAVVTSLFWTFLDVVKVRLQSORP
 SMASLMPSSRLWSLSYTKWKCLLYCNGVLEPLVLCNPNGARCATWFDPRFTGTGND
 FVKIVRHGSTRILWSGLPATLVMTVPATAIYFTAYDQLKFLQGRALTSLDLYAPWVAG

FEATURES
 source

CDS

ALALGCTVTVISPLEIMETKLAQAHVSRYELGACVETAVAGQWMSRLMLGWTALRD
 VPFSALYWNFYNELVSLWNGFRPKDQTSVGMFVAGGIGTVAALVTLPFPDWWKTQ
 VALGAEAVRVNPLHVDSTWLLRRIRAESGTLKGLFAGFLPRIKAAFSCAIMISTYE
 FGKSFQRLNQDRLGG"

BASE COUNT 293 a 496 c 464 g 326 t
 ORIGIN

Query Match 94.6%; Score 1054.4; DB 9; Length 1579;
 Best Local Similarity 97.8%; Pred. No. 2.2e-230;
 Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACAGAACCTCGGGGCATCGAGCCCTCCAGCAAAATGCT 60
 Db 118 CTGAAGCTTCAAGATGGCTGACAGAACCTCGGGGCATCGAGCCCTCCAGCAAAATGCT 177
 QY 61 GGCCTCAGGACCGGGGCTGTGGTTACCTCTCTCTCATGACACCCCTGACGCTGTGAA 120
 Db 178 GGCCTCAGGACCGGGGCTGTGGTTACCTCTCTCTCATGACACCCCTGACGCTGTGAA 237
 QY 121 GGTTCGCTGCAGTCTCAGCGGCCCTCCATGGCAGCGAGCTGATGCTTCTTCCAGACT 180
 Db 238 GGTTCGCTGCAGTCTCAGCGGCCCTCCATGGCAGCGAGCTGATGCTTCTTCCAGACT 297
 QY 181 GTGGAGCCTCTCTATACAAAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTCCCTCT 240
 Db 298 GTGGAGCCTCTCTATACAAAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTCCCTCT 333
 QY 241 GTATTGCAATGCTCTCTGAGGCTCTGTACCTGTGCGCAATGGTCCCGCTGTGCCAC 300
 Db 334 GTATTGCAATGCTCTCTGAGGCTCTGTACCTGTGCGCAATGGTCCCGCTGTGCCAC 393
 QY 301 CTGGTTTCAAGACCTTACCTGCGCACCATGATGCTTCTGTGAAGATCTGTAG 360
 Db 394 CTGGTTTCAAGACCTTACCTGCGCACCATGATGCTTCTGTGAAGATCTGTAG 453
 QY 361 GCACGACCTCAGGACCGCTCTGGAGCGGCTCTCCCGCCACCTGTGTGATGATGCTGCC 420
 Db 454 GCACGAGGACCGGACCGCTCTGGAGCGGCTCTCCCGCCACCTGTGTGATGATGCTGCC 513
 QY 421 AGCTACCGCCATCTACTTCTGCTGATGACCACTGAGGCGCTTCTGTGTGCTGAGC 480
 Db 514 AGCTACCGCCATCTACTTCTGCTGATGACCACTGAGGCGCTTCTGTGTGCTGAGC 573
 QY 481 CTTGACCTCTGACCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 574 CTTGACCTCTGACCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633
 QY 541 GACTGTGATGACCGCTTGTGAGCTTATGCGGACAAAGCTGACGCTGACGCTGCTGCTGCT 600
 Db 634 GACTGTGATGACCGCTTGTGAGCTTATGCGGACAAAGCTGACGCTGACGCTGCTGCTGCT 693
 QY 601 CCGGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 694 CCGGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
 QY 661 GCTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 754 GCTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
 QY 721 CTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGCTTCTGTGGG 780
 Db 814 CTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGCTTCTGTGGG 873
 QY 781 CATGAGCTTCTGTGGCTGGGCTCTCAGGAGCGGTGGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 874 CATGAGCTTCTGTGGCTGGGCTCTCAGGAGCGGTGGCTGCTGCTGCTGCTGCTGCTGCT 933
 QY 841 TCAGCTGTGAAGACCCCAACCGCAGCTGCTCTGGAGCGATGGAGGCTGTGAGAGTGAA 900
 Db 934 TCAGCTGTGAAGACCCCAACCGCAGCTGCTCTGGAGCGATGGAGGCTGTGAGAGTGAA 993
 QY 901 CCCCTGTGATGAGCTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

```

Db 994 CCCCTCATGTGACTCCACCTGGCTGCTGCTGGAGGATCGGGCGAGTCGGGCAC 1053
Qy 961 CAAGGAGCTTTTGCAGCTTCCCTCGATCATCAAGGCTGCCCCCTCTGTGTCAT 1020
Db 1054 CAAGGAGCTTTTGCAGCTTCCCTCGATCATCAAGGCTGCCCCCTCTGTGTCAT 1113
Qy 1021 CATGATCAGCACTATGATTCGGCAAAAGCTTCTCCAGAGGCTGAACCAAGACGGCT 1080
Db 1114 CATGATCAGCACTATGATTCGGCAAAAGCTTCTCCAGAGGCTGAACCAAGACGGCT 1173
Qy 1081 TCTGGCGGCTCAAGAGGGCAAGAGGCAAGGAC 1114
Db 1174 TCTGGCGGCTCAAGAGGGCAAGAGGCAAGGAC 1207

RESULT 8
AR116038 AR116038 1643 bp DNA linear PAT 16-MAY-2001
LOCUS AR116038 Sequence 68 from patent US 6132973.
DEFINITION AR116038
ACCESSION AR116038
VERSION AR116038.1 GI:14096360
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1643)
AUTHORS Lal,P., Hillman,J.L., Bandman,O., Shah,P., Au-Young,J., Yue,H.,
Guegler,K.J. and Corley,N.C.
TITLE Human regulatory molecules
JOURNAL Patent: US 6132973-A 68 17-OCT-2000;
FEATURES
source 1..1643
/organism="unknown"
BASE COUNT 307 a 501 c 476 g 359 t
ORIGIN

Query Match 94.6%; Score 1054.4; DB 6; Length 1643;
Best Local Similarity 97.88; Pred. No. 2.1e-230;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

Qy 1 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCATCAGCCCCCTCCAGCAATGCT 60
Db 192 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGCATCAGCCCCCTCCAGCAATGCT 251
Qy 61 GGCCTCAGGACCCGGGCTGTGTTACTCTCTCTTCATGACACCCCTCGAGCTGTGAA 120
Db 252 GGCCTCAGGACCCGGGCTGTGTTACTCTCTCTTCATGACACCCCTCGAGCTGTGAA 311
Qy 121 GGTTCGCTGAGTCTCAGCGGCCCTCCATGCCAGGAGTGTCTTCTCCAGACT 180
Db 312 GGTTCGCTGAGTCTCAGCGGCCCTCCATGCCAGGAGTGTCTTCTCCAGACT 371
Qy 181 GTGAGCTCTCTCTATACCAAAATTGCCCTCTCTCTCCAATCCACAGGAGTGCCTCT 240
Db 372 GTGAGCTCTCTCTATACCAAAATTGCCCTCTCTCTCCAATCCACAGGAGTGCCTCT 407
Qy 241 GTATTGCAATGTGTCTGAGAGCTGTACTGTGCGCAATGTCGCCGCTGTGCCAC 300
Db 408 GTATTGCAATGTGTCTGAGAGCTGTACTGTGCGCAATGTCGCCGCTGTGCCAC 467
Qy 301 CTGTTTCAAGACCTACCGCTTCACTGGCACCATGATGCTCTCTGTAAGATCTGTAG 360
Db 468 CTGTTTCAAGACCTACCGCTTCACTGGCACCATGATGCTCTCTGTAAGATCTGTAG 527
Qy 361 GCACGAGGACACAGGACCTCTTGAGGCGGCTCCCGCCACCCCTGGTGTACTGTGCC 420
Db 528 GCACGAGGACACAGGACCTCTTGAGGCGGCTCCCGCCACCCCTGGTGTACTGTGCC 587
Qy 421 AGCTACCGGCATCTACTTCACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTCGAGC 480
Db 588 AGCTACCGGCATCTACTTCACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTCGAGC 647
Qy 481 CCTGACCTCTGACCTTACGACCAACCATGTTGGTGGCGGCTGCCCGCTGGGACCTGT 540

```

```

Db 648 CTTGACCTTCTACCTTCTACGACCCCATGTGTGCTGGCGCTGGCGCTTGGGACCGT 707
Qy 541 GACTGTGATCAGCCCCCTTGGAGCTTATGCGGCAAAAGCTGAGGCTCAGCATGTGCTGA 600
Db 708 GACTGTGATCAGCCCCCTTGGAGCTTATGCGGCAAAAGCTGAGGCTCAGCATGTGCTGA 767
Qy 601 CCGGGCTGGGTGCTGTGTTGAACTCAGTGGCTCAGGCTGGCTGGGCTCAGTGG 660
Db 768 CCGGAGCTGGGTGCTGTGTTGAACTCAGTGGCTCAGGCTGGCTGGGCTCAGTGG 827
Qy 661 GCTGGGCTGGGSCCCCACTGCCCTTCGAGATGTCCTTCTCAGCCCTGACTGTTCAA 720
Db 828 GCTGGGCTGGGSCCCCACTGCCCTTCGAGATGTCCTTCTCAGCCCTGACTGTTCAA 887
Qy 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGCAAGGACCAAGCTTCTGTGG 780
Db 888 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGCAAGGACCAAGCTTCTGTGG 947
Qy 781 CATGAGCTTTTGGGCTGGTGGCATCTCAGGACGCTGGCTCAGTGTGCTGCTCAGCTT 840
Db 948 CATGAGCTTTTGGGCTGGTGGCATCTCAGGACGCTGGCTCAGTGTGCTGCTCAGCTT 1007
Qy 841 TGACGTGGTAAAGACCCCAAGCCAGGCTCTCTGGGAGCGATGAGGCTGTGAGAGTGA 900
Db 1008 TGACGTGGTAAAGACCCCAAGCCAGGCTCTCTGGGAGCGATGAGGCTGTGAGAGTGA 1067
Qy 901 CCCCCTGATGTGAGACTCCACCTGGCTGTCTGGGAGGATCCGGCCGAGTCGGGAC 960
Db 1068 CCCCCTGATGTGAGACTCCACCTGGCTGTCTGGGAGGATCCGGCCGAGTCGGGAC 1127
Qy 961 CAAGGAGCTTTTGGGCTGGTGGCATCTCAGGATCATCAAGGCTGCCCTCTCTGTGCCAT 1020
Db 1128 CAAGGAGCTTTTGGGCTGGTGGCATCTCAGGATCATCAAGGCTGCCCTCTCTGTGCCAT 1187
Qy 1021 CATGATCAGCACTATGAGTTCGGCAAAAGCTTCTTCAGAGGCTGAACCAAGACCGCT 1080
Db 1188 CATGATCAGCACTATGAGTTCGGCAAAAGCTTCTTCAGAGGCTGAACCAAGACCGCT 1247
Qy 1081 TCTGGCGGCTCAAGAGGGCAAGGAGGCAAGGAC 1114
Db 1248 TCTGGCGGCTCAAGAGGGCAAGGAGGCAAGGAC 1281

RESULT 9
AK026060 AK026060 1554 bp mRNA linear PRI 29-SEP-2000
LOCUS Homo sapiens cDNA: FLJ22407 fis, clone HRC08407.
DEFINITION AK026060
ACCESSION AK026060
VERSION AK026060.1 GI:10438782
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens primary human renal epithelial cells cDNA to mRNA,
clone lib:HRC clone:HRC08407.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1554)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert

```

sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES		Location/Qualifiers
Source	1..1554	/organism="Homo sapiens"
	/db_xref="taxon:9606"	
	/clone="HRC08407"	
	/cell_type="primary human renal epithelial cells"	
	/clone_lib="HRC"	
CDS	118..1173	/note="cloning vector pME-8SFL3"
	118..1173	/note="unnamed protein product"
	/codon_start=1	
	/protein_id="BAB15341.1"	
	/db_xref="GI:10438783"	
BASE COUNT		285 a 489 c 457 g 323 t
ORIGIN		FGKSFQRINQDRLLG"
Query Match		94.5%; Score 1052.8; DB 9; Length 1554;
Best Local Similarity		97.7%; Pred. No. 5e-230;
Matches 1088; Conservative		0; Mismatches 2; Indels 24; Gaps 1;
QY	1	CTGAAGCTTCAAGTGGCTGACGAGACCTCGGGGATCAGCCCTCCAGCAATGGT 60
DB	105	CTGAAGCTTCAAGTGGCTGACGAGACCTCGGGGATCAGCCCTCCAGCAATGGT 164
QY	61	GGCTTCAGGACCGGGCTGTGTACTCTCTCTTATGACACCCCTGGAGCTGGTAA 120
DB	165	GGCTTCAGGACCGGGCTGTGTACTCTCTCTTATGACACCCCTGGAGCTGGTAA 224
QY	121	GGTTCGCTGAGTCTCAGGGGCTTCCATGGCCAGCGAGTATGCTTCTCCAGACT 180
DB	225	GGTTCGCTGAGTCTCAGGGGCTTCCATGGCCAGCGAGTATGCTTCTCCAGACT 284
QY	181	GTGAGGCTCTCTATACCAATGTGCTCTCTCTCCATCCAGGAGTGGCTCT 240
DB	285	GTGAGGCTCTCTATACCAATGTGCTCTCTCTCCATCCAGGAGTGGCTCT 320
DB	241	GTATTGCAATGGTCTCTGAGCTCTGTACTGTGCCCCAAATGGTGGCTGGCCAC 300
DB	321	GTATTGCAATGGTCTCTGAGCTCTGTACTGTGCCCCAAATGGTGGCTGGCCAC 380
QY	301	CTGGTTTCAAGACCTTACCGCTTCACTGGACACCATGGATGCTTGTGAAGTGTGAG 360
DB	381	CTGGTTTCAAGACCTTACCGCTTCACTGGACACCATGGATGCTTGTGAAGTGTGAG 440
QY	361	GCACGAGGACACGAGACCTCTGAGAGGGCTCTCCCGCCACCTGGTATGACTGTGCC 420
DB	441	GCACGAGGACACGAGACCTCTGAGAGGGCTCTCCCGCCACCTGGTATGACTGTGCC 500
QY	421	AGCTACCGCCATCTACTTCACTGCTATGACCAACTGAAGCCCTTCTGTGTGTCGAGC 480
DB	501	AGCTACCGCCATCTACTTCACTGCTATGACCAACTGAAGCCCTTCTGTGTGTCGAGC 560
QY	481	CCTGACCTCTACCTCTACGACCATGGTGGCTGGCGCTGGCCGCTGGGACCGT 540
DB	561	CCTGACCTCTACCTCTACGACCATGGTGGCTGGCGCTGGCCGCTGGGACCGT 620
QY	541	GACTGTGATCAGCCCCCTGGAGCTTATGCGACAAAGCTCGAGGCTCAGCATGTGCTGA 600
DB	621	GACTGTGATCAGCCCCCTGGAGCTTATGCGACAAAGCTCGAGGCTCAGCATGTGCTGA 680
QY	601	CCGGAGCTGGGTGCTGTGTTTGGAACTGAGTGGCTCAGGGTGGCTGGCTCACTGTG 660

Db	681	CCGGAGCTGGGTGCTGTGTTGAACTCAGTGGCTCAGGGTGGCTCACTGTG 740
QY	661	GCTGGGCTGGGGCCCACTGCCCTTCGAGATGTCCTTCTCAGCCCTGACTGTTCAA 720
Db	741	GCTGGGCTGGGGCCCACTGCCCTTCGAGATGTCCTTCTCAGCCCTGACTGTTCAA 800
QY	721	CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAGAGCCAGACTTCTGTGG 780
Db	801	CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAGAGCCAGACTTCTGTGG 860
QY	781	CATGAGCTTGTGGCTGGTGGATCTCAGGACGCTGCTCAGTGTGCTGACTACCTT 840
Db	861	CATGAGCTTGTGGCTGGTGGATCTCAGGACGCTGCTCAGTGTGCTGACTACCTT 920
QY	841	TGACGTGCTTAAGACCCCAAGCCAGGCTGCTCTGGAGCGATGGAGCTGTGAGAGTGA 900
Db	921	TGACGTGCTTAAGACCCCAAGCCAGGCTGCTCTGGAGCGATGGAGCTGTGAGAGTGA 980
QY	901	CCCCCTGATGTGACTTCCACCTGGCTGCTGCTGGAGGATCCGGCCGAGTCGGGCAC 960
Db	981	CCCCCTGATGTGACTTCCACCTGGCTGCTGCTGGAGGATCCGGCCGAGTCGGGCAC 1040
QY	961	CAAGGACTCTTTTCAGGCTTCTCTCTCGATCATCAAGGCTGCCCTTCTGTGCCAT 1020
Db	1041	CAAGGACTCTTTTCAGGCTTCTCTCTCGATCATCAAGGCTGCCCTTCTGTGCCAT 1100
QY	1021	CATGATCAGCACTATGATTCGGGAAAGCTTCTTCAGAGGCTGAACAGGACCGCT 1080
Db	1101	CATGATCAGCACTATGATTCGGGAAAGCTTCTTCAGAGGCTGAACAGGACCGCT 1160
QY	1081	TCTGGGCGGCTGAAAGGGGCAAGGAGCAAGCAC 1114
Db	1161	TCTGGGCGGCTGAAAGGGGCAAGGAGCAAGCAC 1194
RESULT 10		
AF119864		
ID	AF119864	standard, RNA; HUM; 1402 BP.
XX	AC	AF119864;
XX	XX	AF119864.1
XX	XX	12-MAY-2000 (Rel. 63, Created)
DT	09-MAY-2001	(Rel. 67, Last updated, Version 2)
XX	XX	Homo sapiens PRO2163 mRNA, complete cds.
DE	XX	
XX	XX	
XX	XX	
OS	XX	Homo sapiens (human)
OC	XX	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC	XX	Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX	XX	[1]
RP	1-1402	Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RA	1-1402	"Functional prediction of the coding sequences of 79 new genes deduced by
RT	1-1402	analysis of cDNA clones from human fetal liver";
RT	1-1402	Unpublished.
XX	XX	[2]
RP	1-1402	Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RA	1-1402	Submitted (13-JAN-1999) to the EMBL/GenBank/DBJ databases.
RL	1-1402	Department of Experimental Hematology, Institute of Radiation Medicine,
RL	1-1402	Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
XX	XX	SWISS-PROT; Q9B2J4; CG69_HUMAN.
DR	XX	
XX	XX	Key
XX	XX	Location/Qualifiers


```

FT source
FT 1. .1402
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="FLB8041"
FT /tissue_type="liver"
FT /dev_stage="fetus"
FT 4. .1017
FT /codon_start=1
FT /db_xref="SWISS-PROT:Q9BJJ4"
FT /note="predicted protein of H2163"
FT /product="PR02163"
FT /protein_id="AAF69618.1"
FT /translation="WVAGTGAIVTSLFMTPLDVVKVLRQSFMSAMELPPSRILWSL
FT SYTKMKLLYKNGVLEPLYLCPNGARCATWFDPTFRFTGTMDFVKIVRHEGRTILWSG
FT LPTALVMTVPATAIYFTAYDOLKAFLCGRALTDLYAPWVAGALARLGTVTVISPLELM
FT RTLOACHVYSRELCAVCRTAVAOGGWRSLWLGWPTALRDVPFSALYWNELVKSLL
FT NGLRPKDOTSVMKSVFVAGGISRTVAALVTLTPEVYKTOEVALGAMAEAVNPVLDST
FT WLLRIRAESGTLGFLAGFLPRLIKAAPSCAIIIMSTYEFKSFQRLNDRLLGG"
FT XX
FT SQ Sequence 1402 BP; 261 A; 432 C; 401 G; 308 T; 0 other;

Query Match 90.0%; Score 1002.4; DB 17; Length 1402;
Best Local Similarity 97.6%; Pred. No. 1.6e-218;
Matches 1037; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 53 CAATGTTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGAC 112
DB 1 CAATGTTGGCTCAGGACCGGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGGAC 60

QY 113 GTGGTGAAGGTTGCGCTGAGTCTCAGCGGCTCTCATGGCCAGCAGCTGATGCTCTCC 172
DB 61 GTGGTGAAGGTTGCGCTGAGTCTCAGCGGCTCTCATGGCCAGCAGCTGATGCTCTCC 120

QY 173 TCCAGACTGTGGAGCTCTCTATACCAAAATGCCCTCTCTCTCCAAATCCACAGGGAAG 232
DB 121 TCAGACTGTGGAGCTCTCTATACCAAAATGCCCTCTCTCTCCAAATCCACAGGGAAG 156

QY 233 TGCTCTCTGTATTCGAATGTTCTCTGGAGCTCTGTACCTGTGCCCAATGTCGCCGC 292
DB 157 TGCTCTCTGTATTCGAATGTTCTCTGGAGCTCTGTACCTGTGCCCAATGTCGCCGC 216

QY 293 TGTGCCACCTGTTTCAAGACCTACCGCTTCACTGGCAGCTGATGCTCTGTGAG 352
DB 217 TGTGCCACCTGTTTCAAGACCTACCGCTTCACTGGCAGCTGATGCTCTGTGAG 276

QY 353 ATCGTGAGGACGAGGACACGAGACCTCTGTGAGCGGCTCTCCCGCCACCTGTGTATG 412
DB 277 ATCGTGAGGACGAGGACACGAGACCTCTGTGAGCGGCTCTCCCGCCACCTGTGTATG 336

QY 413 ACTGTGCCAGTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGT 472
DB 337 ACTGTGCCAGTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGT 396

QY 473 GGTGAGCGCTTGAACCTGTGACCTTACGACCATGCTGGCTGGCGGCTGGCGCCGCTG 532
DB 397 GGTGAGCGCTTGAACCTGTGACCTTACGACCATGCTGGCTGGCGGCTGGCGCCGCTG 456

QY 533 GGCACCGTGACTGTGATCAGCGCCCTTGAGCTTTATGCGGCAAAAGCTGCAGGCTCAGCAT 592
DB 457 GGCACCGTGACTGTGATCAGCGCCCTTGAGCTTTATGCGGCAAAAGCTGCAGGCTCAGCAT 516

QY 593 GTGTCTACCGGAGCTGGGTGCTGTGTTCGAATCTGCAAGTGGCTGAGGCTGCTGGCGC 652
DB 517 GTGTCTACCGGAGCTGGGTGCTGTGTTCGAATCTGCAAGTGGCTGAGGCTGCTGGCGC 576

QY 653 TCAGTGTGCTGGGCTGGGCGCCCACTGCGCTTCAGATGTGCCCTTCTCAGCCCTGTAC 712
DB 577 TCAGTGTGCTGGGCTGGGCGCCCACTGCGCTTCAGATGTGCCCTTCTCAGCCCTGTAC 636

QY 713 TGGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGCAAGGACAGACT 772
DB 637 TGGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGCAAGGACAGACT 696
```

```

QY 773 TCTGTGGCATGAGCTTTGTGGCTGTGGCATCTCAGGACGTTGGCTGCAGTGTGTACT 832
DB 697 TCTGTGGCATGAGCTTTGTGGCTGTGGCATCTCAGGACGTTGGCTGCAGTGTGTACT 756
QY 833 CTACCCCTTTGAGCTGTTAAAGACCCAAACCCAGGTCGCTCTGGAGCGATGAGGCTGTG 892
DB 757 CTACCCCTTTGAGCTGTTAAAGACCCAAACCCAGGTCGCTCTGGAGCGATGAGGCTGTG 816
QY 893 AGAGTGAACCCCTTCGATGGACTCCACCTGCTGCTGCTGCGGAGGATCCGGGCGGAG 952
DB 817 AGAGTGAACCCCTTCGATGGACTCCACCTGCTGCTGCTGCGGAGGATCCGGGCGGAG 876
QY 953 TGGGCAACCAAGGAGCTCTTTTGAGGCTTCTCTTCTCGGATCATCAAGGCTGCCCTCTCC 1012
DB 877 TGGGCAACCAAGGAGCTCTTTTGAGGCTTCTCTTCTCGGATCATCAAGGCTGCCCTCTCC 936
QY 1013 TGTGCCATCATGATCAGACCTATGAGTTTCGCAAAAGCTTCTTCCAGAGGCTGAACGAG 1072
DB 937 TGTGCCATCATGATCAGACCTATGAGTTTCGCAAAAGCTTCTTCCAGAGGCTGAACGAG 996
QY 1073 GACCGCTTCTGGGCGGCTGAAAGGGCGCAAGGAGCAAGGAC 1114
DB 997 GACCGCTTCTGGGCGGCTGAAAGGGCGCAAGGAGCAAGGAC 1038

RESULT 11
AC098934 173434 bp DNA linear HTG 06-NOV-2001
LOCUS Homo sapiens chromosome 1 clone RP11-480112, WORKING DRAFT
DEFINITION SEQUENCE, 6 unordered pieces.
ACCESSION AC098934 AL390963
VERSION AC098934.1 GI:16751904
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
JOURNAL Direct Submission
AUTHORS Unpublished
TITLE 2 (bases 1 to 173434)
JOURNAL Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
AUTHORS Direct Submission
TITLE Submitted (06-NOV-2001) Genome Center, University of Washington,
JOURNAL Box 352145, Seattle, WA 98195, USA
COMMENT On Nov 6, 2001 this sequence version replaced gi:13157584.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgctgs@u.washington.edu
Drafting Center: SC
----- Project Information
Center project name: chr-1
Center clone name: RP11-480112 (sc0719)
----- Summary Statistics
Sequencing vector: plasmid; 32% of reads
Sequencing method: 108752; 68% of reads
Chemistry: Dye-terminator ET; 89% of reads
Chemistry: Dye-terminator Big Dye; 11% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 170284 bases at least Q40
Consensus quality: 171568 bases at least Q30
Consensus quality: 172298 bases at least Q20
Insert size: 172934; sum-of-contigs
Quality coverage: 8.7x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
```


Correll, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsom, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsegad, H.,
Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogah, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 128802)

Worley K.C.

Direct Submission

Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 128802)

Worley K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Jul 3, 2002 this sequence version replaced gi:21239891.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GMFG

Center clone name: CH230-11K3

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 108533 bases at least Q40

Consensus quality: 111720 bases at least Q30

Consensus quality: 113548 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently
consists of 29 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1434: contig of 1434 bp in length

1435 1534: gap of unknown length

1535 3218: contig of 1684 bp in length

3219 3318: gap of unknown length

3319 4531: contig of 1213 bp in length

4532 4531: gap of unknown length

4632 5802: contig of 1171 bp in length

5803 5902: gap of unknown length

5903 7060: contig of 1158 bp in length

7061 7160: gap of unknown length

7161 8514: contig of 1354 bp in length

8515 8614: gap of unknown length

8615 9672: contig of 1058 bp in length

9673 9772: gap of unknown length

9773 12180: contig of 2408 bp in length

12181 12280: gap of unknown length

12281 13463: contig of 1183 bp in length

13464 13563: gap of unknown length

13564 16106: contig of 2543 bp in length

16107 16206: gap of unknown length

16207 18884: contig of 2678 bp in length

18885 18984: gap of unknown length

18985 20532: contig of 1548 bp in length

20533 20632: gap of unknown length

20633 23210: contig of 2578 bp in length

23211 23310: gap of unknown length

23311 25515: contig of 2205 bp in length

25516 25615: gap of unknown length

25616 27411: contig of 1796 bp in length

27412 27511: gap of unknown length

27512 29449: contig of 1938 bp in length

29450 32307: contig of 2758 bp in length

32308 32407: gap of unknown length

32408 35335: contig of 3128 bp in length

35336 35636: gap of unknown length

35637 39153: contig of 3518 bp in length

39154 39253: gap of unknown length

39254 43671: contig of 4418 bp in length

43672 43771: gap of unknown length

43772 47082: contig of 3311 bp in length

47083 47182: gap of unknown length

47183 52198: contig of 5016 bp in length

52199 52988: gap of unknown length

52989 57787: contig of 3989 bp in length

57788 57887: gap of unknown length

57888 60500: contig of 4113 bp in length

60501 60600: gap of unknown length

60601 66994: contig of 6394 bp in length

66995 67094: gap of unknown length

67095 73183: contig of 6089 bp in length

73184 73283: gap of unknown length

73284 80968: contig of 7685 bp in length

80969 81068: gap of unknown length

81069 98060: contig of 16992 bp in length

98061 98160: gap of unknown length

98161 128802: contig of 30642 bp in length.

FEATURES

source

1..128802
/organism="Rattus norvegicus"
/db_xref="taxon:101116"
/clone="CH230-11K3"

BASE COUNT 21721 a 38987 c 24403 g 38221 t 5470 others

ORIGIN

Query Match 50.3%; Score 560.8; DB 2; Length 128802;
Best Local Similarity 79.0%; Pred. No. 7.9e-118;
Matches 719; Conservative 0; Mismatches 182; Indels 9; Gaps 4;

Qy 2 TGAAGCTTCAAGATGGCTGACAGGACCCCTCGGGGCATCAGCCCTCCAGCAATGGTG 61

Db 78069 TGAATTCACAGATGGTTGACAGGATCTGGTGGCATTAGACCCCTTTTACCAATGGTG 78010

Qy 62 GCCTCAGGCACGGGGCTGGGTACCTCTCTCTCATGACACCCCTGAGACCTGGTGAAG 121

Db 78009 GCCTCAGGAACCTGGGGTTGGTGGTGCATCTCCCTTTCATGACACCCCTGGTGTGAAG 77950

Qy 122 GTTCGCTCGAGTCTCA-----CGGCCCTCCATGCGCCAGGAGCTGATGCTTCTCTCA 176

Db	77949	GTCCGCCTTCAGTCTCAAGACCTGTCCCTCGATAGATAGCAAAATGACAAACACCTTCCA	77899
QY	177	GACTGTGGAGCCTCTCCTATACAAATGGCCCTCTCTCTCCAATCCAAGGAAGTGCC	236
Db	77889	GATTCGTGGAGTCTTCTCTACACCAATCACCCCTCTCTCTCAATCCCCAGTGACATGCT	77830
QY	237	TCCTGTAATTGCAATGGTGTCCTGGAGCCTCTGTACCTGTGGCCAAATGGTGGCCGCTGTG	296
Db	77829	TCTCTGTACTGCAATGGTGTCTGGAGCCCTGTACCTGTGCCCAAATGGTACTCGTGCA	77770
QY	297	CCACCTGCTTTCAAGACCCCTACCCCTTCACTGTGCACCAATGATGCCCTTGTGAAGATCG	356
Db	77769	CCACCTGCTTTCAGATCTCTACACAGTTCCTACTAGACCTTATAGTCCCTTGTGGAAGCTG	77710
QY	357	TGAGGCACGAGGCCACGAGACCCCTCTGGAGCGCCTCCCGCCACCCCTGTGATGACTG	416
Db	77709	TGAGGCATAGGGCACGAGACCCCTGTAGAGACGCTCCACGCCACCGTGGTATGACTG	77650
QY	417	TGCCAGCTACCGCATCTACTTCACTGCGCTATGACCAACTGAAGGCCCTTCTCTGTGGTC	476
D	7649	TGTCTGCTACCATCAGTACTTCTACAGCTTATGAGCAGCTCAAGTCCCTCTATGTGGTC	77590
QY	477	GAGCCCTCAGCTCTGACCTCTAGCACCCATGGTGGTGGCGCGCTGGCCCTGGGCA	536
Db	77589	AGTTCTTAACTCTGACCTCTTACGACCCCATAGTAGCTGGTCCCTGGTGTAATGGGCA	77530
QY	537	CGTGACTGTGATAGCCCCCTGGAGCTTATGGGGACAAAGCTGCAGGCTCAGCATGTGT	596
Db	77529	TTGTGCACAGTGGTGAGCCCTTGGAGCTTGTGGGATCAAGCTGCAGGCTCAGCATGT	77470
QY	597	CGTACGGGAGCTGGGTGCTGTGTTCGAATCAGTGGCTCAGGCTGGCTGGCGCTCAC	656
Db	77469	CATACAGTGAATCGGCTACTGTGTTCAGCTGTGGTAGCTCAGAGTGGCTGGCGCTCAC	77410
QY	657	TGTGCTGGGCTGGGGCCCACTGCCCTTCAGATGTGCCCTCTTCAGC-CCTGTACTGG	715
Db	77409	TGTGCTGGGCTGGGCTCTCAGCTCCTCAAGATGAACTTCTCAGCTCTTGTACTGG	77350
QY	716	TTCAACTATGAGCTGTGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAGGACCCAGATCT	775
Db	77349	GTCAACTACGAGCTGGTAAAGAGCTGGCTGAGTGGACT-AAGACAAAGAACAGACATCT	77291
QY	776	GTGGGCATGAGCTTTGTGCTGGTGGGATCTCAGGAGACGGTGGCTCAGTGTGACTCTA	835
Db	77290	GTGAGCGGAGCTTTGTGCTGGTGGCACTCTCAGGAATGGTGGCTGC-CAGTTACCTTA	77233
QY	836	CCCTTTGACGTGTGAAGACCCCAACCGCAGGTGGCTCTCTGGGAGCGATGGAGGCTGTGAGA	895
Db	77232	CTTTTCCATGTGGTGAAGACACAGCCCAAAATGTCACTGGGAGCAGTGGAGGCTGTGAGA	77173
QY	896	GTGAACCCCC	905
Db	77172	GGTTTCTCTCC	77163

```

RESULT 13
AC119489_3/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AC119489 Accession AC119489
Fragment Name Begin End
AC119489_0 1 110000
AC119489_1 100001 210000
AC119489_2 200001 310000
AC119489_3 300001 355252
Continuation (4 of 4) of AC119489 from base 300001 (AC119489 Rattus norvegicus clone CH2
Query Match 50.2%; Score 558.8; DB 2; Length 55252;
Best Local Similarity 74.9%; Fred. No. 2.4e-117;
Matches 321; Conservative 0; Mismatches 182; Indels 93; Gaps 6;
Qy 2 TGAAGCTTCAGATPGCTGACACGAGACCTCGGGGCATCAGCCCTCCAGCAAAATGGT 61
Db 19871 TGAAACTCCAAGATGGATGACACAGANTCTCGGGGCATTAGTCCCTTCAGCAAAATGTG 19812

```

RESULT 14

AC111356

LOCUS

DEFINITION

AC111356

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC111356 156175 bp DNA linear HTG 13-JUL-2002
Rattus norvegicus clone CH230-161G12, *** SEQUENCING IN PROGRESS
***, 66 unordered pieces.

AC111356.2 GI:21735881

HTG; HTGS PHASE1.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 156175)

Murny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amarutunge,H.C., Are,J.R., Ayelle,M., Banks,T.,
Barbacia,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,K., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elnaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J.J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,W., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nckenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Tang,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Thomas,K., Thomas,S.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Wang,Q.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williams,A., Wleciyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 156175)

Worley,K.C.

Direct Submission

Submitted (19-FEB-2002)

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 156175)

Worley,K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18701120.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GMCE

Center clone name: CH230-161G12

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 99973 bases at least Q40

Consensus quality: 106615 bases at least Q30

Consensus quality: 111100 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 66 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will

* be preserved.

* 1 1025: contig of 1025 bp in length

* 1026 1125: gap of unknown length

* 1126 2218: contig of 1093 bp in length

* 2219 2318: gap of unknown length

* 2319 3454: contig of 1136 bp in length

* 3455 3554: gap of unknown length

* 3555 4690: contig of 1136 bp in length

* 4691 4790: gap of unknown length

* 4791 5955: contig of 1165 bp in length

* 5956 6056: gap of unknown length

* 6056 7430: contig of 1384 bp in length

* 7430 7539: gap of unknown length

* 7540 8707: contig of 1168 bp in length

* 8708 8807: gap of unknown length

* 8808 10033: contig of 1226 bp in length

* 10034 10133: gap of unknown length

* 10134 11778: contig of 1645 bp in length

* 11779 13633: contig of 1755 bp in length

* 13634 13733: gap of unknown length

* 13734 14947: contig of 1214 bp in length

* 14948 15047: gap of unknown length

* 15048 16569: contig of 1522 bp in length

* 16570 16669: gap of unknown length

* 16670 17786: contig of 1117 bp in length

* 17787 17886: gap of unknown length

* 17887 19738: contig of 1852 bp in length

* 19739 19838: gap of unknown length

* 19839 21537: contig of 1694 bp in length

* 21538 22633: gap of unknown length

* 22633 22715: contig of 1083 bp in length

* 22716 22815: gap of unknown length

* 22816 24052: contig of 1237 bp in length

* 24053 24152: gap of unknown length

* 24153 25250: contig of 1098 bp in length

* 25251 25944: contig of 1594 bp in length

* 25945 27044: gap of unknown length

* 27045 28474: contig of 1430 bp in length

* 28475 28574: gap of unknown length

* 28575 29925: contig of 1351 bp in length

* 29926 30025: gap of unknown length

* 30026 31402: contig of 1377 bp in length

* 31403 31502: gap of unknown length

* 31503 33645: contig of 2143 bp in length

* 33646 33745: gap of unknown length

* 33746 35290: contig of 1545 bp in length

* 35291 35390: gap of unknown length

* 35391 36672: contig of 1282 bp in length

* 36673 36772: gap of unknown length

Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homs, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, B., Korvan, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pul, L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojudo, I., Rolfe, W., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sison, I.,
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 195701)
 Worley, K.C.
 Direct Submission
 Submitted (13-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 195701)
 Worley, K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 17, 2002 this sequence version replaced gi:21397217.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GYVK
 Center clone name: CH230-75J9
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 143647 bases at least Q40
 Consensus quality: 149058 bases at least Q30
 Consensus quality: 153360 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 60 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1426: contig of 1426 bp in length
 * 1427: gap of unknown length
 * 1526: contig of 1386 bp in length
 * 1527: gap of unknown length
 * 2912: contig of 1386 bp in length
 * 3012: gap of unknown length
 * 3013: contig of 1139 bp in length
 * 3014: gap of unknown length
 * 4251: contig of 1378 bp in length
 * 4252: gap of unknown length
 * 5630: contig of 1197 bp in length
 * 5729: gap of unknown length
 * 5730: contig of 1197 bp in length
 * 6927: gap of unknown length
 * 7026: contig of 1137 bp in length
 * 7027: gap of unknown length
 * 8163: contig of 1319 bp in length
 * 8164: gap of unknown length
 * 8263: contig of 1319 bp in length
 * 8264: gap of unknown length
 * 9582: contig of 1707 bp in length
 * 9583: gap of unknown length
 * 9584: contig of 1565 bp in length
 * 11390: gap of unknown length
 * 11391: contig of 1565 bp in length
 * 13054: gap of unknown length
 * 13055: contig of 1256 bp in length
 * 13155: gap of unknown length
 * 14410: contig of 1080 bp in length
 * 14411: gap of unknown length
 * 15590: contig of 1292 bp in length
 * 15591: gap of unknown length
 * 15982: contig of 1835 bp in length
 * 15983: gap of unknown length
 * 17083: contig of 1713 bp in length
 * 18917: gap of unknown length
 * 19018: contig of 1309 bp in length
 * 20731: gap of unknown length
 * 20730: contig of 1309 bp in length
 * 22139: gap of unknown length
 * 22140: contig of 1367 bp in length
 * 22240: gap of unknown length
 * 23607: contig of 1537 bp in length
 * 23706: gap of unknown length
 * 25243: contig of 1250 bp in length
 * 25244: gap of unknown length
 * 25943: contig of 1721 bp in length
 * 26594: gap of unknown length
 * 26595: contig of 2921 bp in length
 * 28415: gap of unknown length
 * 28514: contig of 1996 bp in length
 * 31535: gap of unknown length
 * 31536: contig of 2695 bp in length
 * 33531: gap of unknown length
 * 33532: contig of 2695 bp in length
 * 33533: gap of unknown length
 * 36327: contig of 3647 bp in length
 * 36328: gap of unknown length
 * 40074: contig of 1958 bp in length
 * 40173: gap of unknown length
 * 42131: contig of 2222 bp in length
 * 42132: gap of unknown length
 * 42232: contig of 1630 bp in length
 * 44454: gap of unknown length
 * 44554: contig of 2101 bp in length
 * 46184: gap of unknown length
 * 46284: contig of 3760 bp in length
 * 48385: gap of unknown length
 * 48485: contig of 2343 bp in length
 * 52245: gap of unknown length
 * 52345: contig of 2424 bp in length
 * 54688: gap of unknown length
 * 54788: contig of 1887 bp in length
 * 57212: gap of unknown length
 * 57312: contig of 2227 bp in length
 * 59199: gap of unknown length
 * 59299: contig of 2512 bp in length
 * 61526: gap of unknown length
 * 61626: contig of 2757 bp in length
 * 64138: gap of unknown length
 * 64237: contig of 2388 bp in length
 * 66995: gap of unknown length
 * 67095: contig of 2388 bp in length
 * 69582: gap of unknown length

Query Match		49.6%; Score 552.8; DB 2; Length 195701;
Best Local Similarity		83.4%; Pred. No. 5.1e-116;
Matches 664; Conservative		0; Mismatches 127; Indels 5; Gaps 3;
Qy	302	TGTTTTCAGACCTTACCCGCTTCTAGTGGACCATGGATGGCTTCGTGAAGATCGTGAGG 361
Db	72766	TGTTTTCAGATCTTACACGGTTTCACTGGCACCTTGGATGGCTTCGTGAAGATTTGAGG 72825
Qy	362	CAGGAGGACACGAGACCTCTGAGCGGCTCCCGCACCTGGTGTGATGACTGTGCCA 421
Db	2826	CAGGAGGACACGAGACCTCTGAGCGGCTCCCGCACCTGGTGTGATGACTGTGCCA 72885
Qy	422	GCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGCTCGAGCC 481
Db	72886	GCCACTGCCATCTACTTCACTGCTATGACCAACTGAAGGCTTCCTGTGTGCTCGAGCC 72945
Qy	482	CTGACCTCTGACCTTACGACCACTGGCTGGCTGGCGGCTGGCCGCTGGGCAACGCTG 541
Db	72946	TTGACCTCTGACCTTATGACCACTGGCTGGCTGGCGGCTGGCCGCTGGGCAACGCTG 73005
Qy	542	ACTGTGATCAGCCCTCTGGAGCTTATGCGGACAAAGCTGCAAGGCTCAGCATGTGCTAC 601
Db	73006	ACGGTGGTCAAGCTTGGAGCTTGTGTGGAACAAGCTGCAAGGCTCAGCATGTGCTAC 73065
Qy	602	CGGAGCTGGGCTGTGTTCGAATGCAAGTGGCTCAGGCTGGCTGGCGCTCAGCTGTGG 661
Db	73066	CGGAACTGGCTGTGTTCGAATGCAAGTGGCTCAGGCTGGCTGGCGCTCAGCTGTGG 73125
Qy	662	CTGGGCTGGGCCCCACTGGCCCTTCAGATGGCCCTTCTCAGCCCTGTACTGTTTCAAC 721
Db	73126	CTGGGCTGGGCCCCACTGGCCCTTCAGATGGCCCTTCTCAGCCCTGTACTGTTTCAAC 73185
Qy	722	TATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGGC 781
Db	73186	TATGAGCTGGTGAAGAGCTGGCTCAATGGGCGGAGGACCAAGCTTCTGTGGC 73245

Qy	782	ATGAGCTTTGTGCTGGTGG-CATCTCAGGACGGTGGCTGCAGTGTCTGACTCTACCCCTT 840
Db	73246	ATCAGCTTTGTGCTGGTGGACTTCTCAGGAATGGTGGCTGCCACCCCTTACCCCTACCCCTT 73305
Qy	841	TGAGTGGTAAAGACCCAAACCCAGGTCGCTCTGGAGCGATGGAGGCTGTGAGAGTGAA 900
Db	73306	GGATGTGGTGAAGACACACAGCGCAAAATGTCACTGGGAGCAGTGAAGGCGGTGAGAATGAA 73365
Qy	901	CCCCCTGCATGTGGACTCCACTGGCTGCTGCTGGAGGATCCGGGCGGAGTCGGGCAC 960
Db	73366	GCACCAAGAGCTG-ATCCACTGGCTTCTCAGAGAAATCCAGGCTGAATCTGGCAC 73424
Qy	961	CAAGGGACTCTTTGCAAGGCTTCCTTCCTCGGATCATCAAGGCTGCCCTCTCTGTGCCAT 1020
Db	73425	CAGGGCACTCTTTGCAAGGCTTCCTTCCTCCCGAGG---ATCAAGGCTGGCCTTCTCTGTGCCAT 73481
Qy	1021	CATGATCAGACCTATGAGTTTCGGCAAAAGCTTCTTTCAGAGGCTGAACCAAGGACGGCT 1080
Db	73482	CATGATCAGACCTTACGAGTTTGGCAAAAGCTTCTTACACAGACTCAACACAGGAGCAGCC 73541
Qy	1081	TCTGGGCGGCTGAAG 1096
Db	73542	TCTGGGCGGCTGAAG 73557

Search completed: January 22, 2003, 23:15:31
Job time : 3647.46 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:31:07 ; Search time 255.466 Seconds
(without alignments)
9820.189 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagctcaagatgctg.....agggcaaggaggcaaggac 1114

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125995159 residues
To. number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1114	100.0	1114	24	AAI71019 Human mitochondria
2	1070	98.1	1662	21	AAC76992 Human ORF2547
3	1056	94.8	1560	21	AAC90452 Human uncoupling p
4	1054.4	94.6	1546	24	AAI71020 Human mitochondria
5	472	42.4	704	24	ABI99647 Mouse ischaemic co
6	472	42.4	2170	24	ABK34735 Human cDNA for nov
7	448.4	40.3	758	24	ABL01419 Murine apoptosis r
8	242.2	21.7	6712	21	AC76691 Human ORF2246
9	185.8	16.7	453	24	ABL81508 Human ovarian canc

c	10	170.2	15.3	512	24	ABK62422	Rat sequence diffe
	11	170	15.3	537	21	AAA06606	Human immunogenic
	12	170	15.3	537	22	AAS63815	Human prostate cDN
	13	170	15.3	537	22	AAH93722	Human prostate-spe
	14	170	15.3	537	22	AAH85036	Human prostate-spe
	15	170	15.3	537	22	AAH02787	Prostate tumour an
	16	170	15.3	537	24	ABL95186	Human CGI-69 cDNA
	17	154	13.8	275	22	AAI00555	Human reproductive
	18	154	13.8	275	23	ABK72077	Human cDNA encodin
	19	154	13.8	275	24	ABK91669	cDNA encoding nove
	20	151.4	13.6	1258	24	ABL89865	Human polynucleoti
	21	142.6	12.8	397	21	AAH30252	Human colon cancer
	22	130.4	11.7	1449	23	ABL04893	Drosophila melanog
	23	130.4	11.7	3928	23	ABL04892	Drosophila melanog
	24	122	11.0	422	22	AAK53810	Murine transport a
c	25	112.4	10.1	1445	23	ABL24583	Drosophila melanog
	26	93.4	8.4	452	24	ABN94635	Gene #1133 used to
	27	79.6	7.1	3645	23	ABL24582	Drosophila melanog
	28	78.2	7.0	1529	23	ABL24581	Drosophila melanog
	29	77.4	6.9	2516	24	ABL90270	Human polynucleoti
	30	74.8	6.7	452	22	ABK59262	Human foetal liver
	31	74.8	6.7	452	22	AAK07474	Human brain expres
	32	74.8	6.7	452	22	AAK33259	Human bone marrow
	33	74.8	6.7	452	22	AAI39053	Probe #7739 used t
c	34	74.8	6.7	452	24	ABS08091	Human genome-deriv
	35	74.8	6.7	496	22	ABA59322	Human foetal liver
	36	74.8	6.7	496	22	AAK07540	Human brain expres
	37	74.8	6.7	496	22	AAK33336	Human bone marrow
	38	74.8	6.7	496	22	AAI39123	Probe #7809 used t
	39	74.8	6.7	496	24	ABS08171	Human genome-deriv
c	40	63.8	5.7	367	24	ABL01420	Murine apoptosis r
	41	61	5.5	997	21	AAZ61628	cDNA encoding muri
	42	61	5.5	997	22	AAC99561	Skin cell cDNA, SE
	43	61	5.5	997	24	ABL34713	Murine cDNA isolat
	44	61	5.5	1808	22	AAC99806	Skin cell cDNA, SE
	45	61	5.5	1808	24	ABL34958	Murine cDNA isolat

ALIGNMENTS

RESULT 1	AAI71019	standard; cDNA; 1114 BP.
ID	AAI71019	standard; cDNA; 1114 BP.
AC	AAI71019;	
XX	AAI71019;	
DT	18-MAR-2002	(first entry)
DE	Human mitochondrial carrier protein CGI-69L cDNA.	
KW	CGI-69L; mitochondrial carrier protein; human; metabolic disease; obesity; cachexia; tumour; cancer; infection; immunomodulator;	
KW	antitumour; virucide; antibacterial; anorectic; antidiabetic;	
KW	brown adipose tissue; diagnosis; gene therapy; splice variant; ss.	
OS	Homo sapiens.	
EH	Key	Location/Qualifiers
FT	CDS	14..1093
FT		/*tag= a
PN	WO200198355-A2.	
XX	27-DEC-2001.	
XX	22-JUN-2001; 2001WO-US20117.	
XX	22-JUN-2000; 2000US-213307P.	
PA	(GETH) GENENTECH INC.	
PA	(CURA-) CURAGEN CORP.	
XX		

PI Lewin D, Adams SH, Yu XX;
 XX WPI; 2002-114569/15.
 DR P-PSDB; AAM50569.
 XX CGI-69 polypeptides and polynucleotides useful for treating metabolic
 PT disorders including cachexia, obesity, diabetes and cancers -
 XX Claim 1(a); Page 9-10; 107pp; English.
 XX The present sequence is that of cDNA encoding human CGI-69L (see
 CC AAM50569), a novel splice variant of human CGI-69. Various CGI-69
 CC clones were isolated from human liver upon PCR amplification and
 CC cloning. Numerous clones diverged from the previously known CGI-69
 CC sequence in that they encoded an 8-amino acid insert preceded by a
 CC W64L change, and were termed the 'long version' isoform or CGI-69L.
 CC The invention relates to a novel characterization of CGI-69 as a
 CC mitochondrial carrier protein, the discovery of CGI-69L, and the
 CC discovery of the alteration of the mitochondrial membrane potential
 CC through overexpression of carboxyl-FLAG-tagged CGI-69. CGI-69
 CC (including splice variant) nucleic acids and proteins are useful
 CC for diagnosing and treating metabolic diseases in humans, including
 CC obesity, cachexia and diabetes. Evidence for an important function
 CC for CGI-69 in modifying mitochondrial membrane potential in brown
 CC adipose tissue (BAT) is presented. The mouse orthologue of CGI-69
 CC is up-regulated in cold-treated BAT. CGI-69 may be involved in
 CC cellular thermogenic uncoupling and, therefore, may be used to
 CC diagnose and treat specific perturbations in metabolic pathways.
 CC Altering the expression of CGI-69 through gene therapy provides a
 CC means of treating metabolic diseases, such as obesity or cachexia,
 CC or of increasing or decreasing body weight. Decreasing the
 CC activity of CGI-69 can be used to treat cachexia, tumours, cancers,
 CC viral infections and bacterial infections, while increasing its
 CC activity can be used to treat obesity, tumours, cancers, viral
 CC infections and bacterial infections (all claimed). Methods are
 CC also claimed for determining whether a compound up-regulates or
 CC down-regulates expression of a CGI-69 gene, and of screening for
 CC a mutation in the CGI-69 gene.
 XX Sequence 1114 BP; 194 A; 349 C; 332 G; 239 T; 0 other;

Query Match 100.0%; Score 1114; DB 24; Length 1114;
 Best Local Similarity 100.0%; Pred. No. 4.3e-264;
 Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAAGCTTCAAGATGGCTGACGAGGACCTGCGGGCATCAGCCCTCCAGCAAAATGGT 60
 DB 1 CTGAAGCTTCAAGATGGCTGACGAGGACCTGCGGGCATCAGCCCTCCAGCAAAATGGT 60
 QY 61 GGCCTCAGGACCGGGCTGTGGTACCTCTCTTCTATGACACCCCTGGAGCTGGTGA 120
 DB 61 GGCCTCAGGACCGGGCTGTGGTACCTCTCTTCTATGACACCCCTGGAGCTGGTGA 120
 QY 121 GGTTCGCTGAGTCTCAGCGGCTTCCATGCGCAGGAGTATGCTTCTCCAGACT 180
 DB 121 GGTTCGCTGAGTCTCAGCGGCTTCCATGCGCAGGAGTATGCTTCTCCAGACT 180
 QY 181 GTGGAGCTCTCTATACCAAAATGGCTCTCTCTCCAAATCCACAGGAGTGGCTCT 240
 DB 181 GTGGAGCTCTCTATACCAAAATGGCTCTCTCTCCAAATCCACAGGAGTGGCTCT 240
 QY 241 GTATTGCAATGGTCTCTGAGGCTCTGTACCTGTGCCCAATGGTGGCGCTGGCCAC 300
 DB 241 GTATTGCAATGGTCTCTGAGGCTCTGTACCTGTGCCCAATGGTGGCGCTGGCCAC 300
 QY 301 CTGGTTTCAGACCTACCGCTCTCACTGCGACCATGATGCTTCTGGAAGATCTGTAG 360
 DB 301 CTGGTTTCAGACCTACCGCTCTCACTGCGACCATGATGCTTCTGGAAGATCTGTAG 360
 QY 361 GCACGAGGACACGAGCCCTCTGGAGCGGCTCTCCCGCCACCCCTGGTGTGATGTGCC 420
 DB 361 GCACGAGGACACGAGCCCTCTGGAGCGGCTCTCCCGCCACCCCTGGTGTGATGTGCC 420

QY 421 AGCTACCGCATCTACTTACATGCTATGACCACTCAAGGCTTCTGTGTGTCGAGC 480
 DB 421 AGCTACCGCATCTACTTACATGCTATGACCACTCAAGGCTTCTGTGTGTCGAGC 480
 QY 481 COTGACCTCTGACTCTACGACACCATGGTGGCTGGCGGCTGGCCCGCTGGGCAACCGT 540
 DB 481 COTGACCTCTGACTCTACGACACCATGGTGGCTGGCGGCTGGCCCGCTGGGCAACCGT 540
 QY 541 GACTGTGATCAGCCCCCTCGAGCTTATGCGGACAAAGCTGCAAGCTCAAGCATGTGTGTA 600
 DB 541 GACTGTGATCAGCCCCCTCGAGCTTATGCGGACAAAGCTGCAAGCTCAAGCATGTGTGTA 600
 QY 601 CCGGAGCTGGTGGCTGTGGTGAAGCTGAGTGGCTGAGTGGCTGGCGCTGAGTGG 660
 DB 601 CCGGAGCTGGTGGCTGTGGTGAAGCTGAGTGGCTGAGTGGCTGGCGCTGAGTGG 660
 QY 661 GCTGGCTGGGGGCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAA 720
 DB 661 GCTGGCTGGGGGCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAA 720
 QY 721 CTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCCGAGACCAAGCTTGTGTGG 780
 DB 721 CTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCCGAGACCAAGCTTGTGTGG 780
 QY 781 CATGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGAGTGGCTGACTTACCTT 840
 DB 781 CATGAGCTTTGTGGCTGGTGGCATCTCAGGAGCGTGGCTGAGTGGCTGACTTACCTT 840
 QY 841 TGACGTGTGTAAGACCCCAAGCCAGCTGCTCTGGGAGCGATGGAGGCTGTGAGAGTGA 900
 DB 841 TGACGTGTGTAAGACCCCAAGCCAGCTGCTCTGGGAGCGATGGAGGCTGTGAGAGTGA 900
 QY 901 CCCCTGATGTGAGCTTCACTGGCTGCTGCTGGGAGGATCGGGCCGAGTGGGGAC 960
 DB 901 CCCCTGATGTGAGCTTCACTGGCTGCTGCTGGGAGGATCGGGCCGAGTGGGGAC 960
 QY 961 CAAGGAGCTCTTTGAGGCTTCTCTCGGATCATCAAGGCTGCCCCCTCTGTGGCAT 1020
 DB 961 CAAGGAGCTCTTTGAGGCTTCTCTCGGATCATCAAGGCTGCCCCCTCTGTGGCAT 1020
 QY 1021 CATGATCAGACCTATGAGTTGGCAAAAGCTTCTTCCAGAGCTGAACACGAGCCGGCT 1080
 DB 1021 CATGATCAGACCTATGAGTTGGCAAAAGCTTCTTCCAGAGCTGAACACGAGCCGGCT 1080
 QY 1081 TCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
 DB 1081 TCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
 RESULT 2
 AAC76992
 ID AAC76992 standard; cDNA; 1662 BP.
 XX AAC76992;
 AC AAC76992;
 DT 08-FEB-2001 (first entry)
 XX Human ORF2547 polynucleotide sequence SEQ ID NO:5093.
 Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
 vulnerable; antipapillary; antipapillary; antipapillary; antipapillary;
 anticonvulsant; osteopathic; antipapillary; antipapillary; antipapillary;
 immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 hypotensive; dermatological; immunosuppressive; antineoplastic;
 antiviral; antibacterial; antifungal; antineoplastic; antineoplastic;
 antineoplastic; cancer; proliferative disorder; hypertension;
 neurodegenerative disorder; osteoarthritis; graft vs host disease;
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 cholesterol ester storage; systemic lupus erythematosus; infection;
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 bone damage; cartilage damage; antiinflammatory disease; coagulation;
 thrombosis; contraceptive; ss.

XX Homo sapiens.
OS WO200058473-A2.
PN 05-OCT-2000.
PD 31-MAR-2000; 2000WO-US08621.
PF 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;
XX WPI; 2000-602362/57.
DR -PSDB; AAB42783.
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX Claim 5; Page 4281-4282; 5507pp; English.
XX AAC74446 to AAC77606 encode the proteins given in BAB40237 to BAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antiproliferative; antiparkinsonian; neurotrophic; immunoprotective;
CC osteoprotective; anticonvulsant; antiarthritic; immunosuppressive;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy.
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
S_q Sequence 1662 BP; 336 A; 516 C; 468 G; 341 T; 1 other;
Query Match 96.1%; Score 1070; DB 21; Length 1662;
Best Local Similarity 97.5%; Pred. No. 3.2e-253;
Matches 1104; Conservative 0; Mismatches 10; Indels 19; Gaps 1;
QY 1 CTGAGCTTCAAGATGGGTGACAGGACCCCTGGGGCATCAGCCCCCTCCAGCAAAATGGT 60
DB 123 CTGAGCTTCAAGATGGGTGACAGGACCCCTGGGGCATCAGCCCCCTCCAGCAAAATGGT 182
QY 61 GGCTCAGGACCCGGGGTGGTGTACCTCTCTTCATGACACCCCTGACGTGGTAA 120
DB 183 GGCTCAGGACCCGGGGTGGTGTACCTCTCTTCATGACACCCCTGACGTGGTAA 242
QY 121 GGTTCGCTCGAGTCTCAGCGGCCCTCCATGGCCAGCGGTGATGCTTCCTCCAGACT 180
DB 243 GGTTCGCTCGAGTCTCAGCGGCCCTCCATGGCCAGCGGTGATGCTTCCTCCAGACT 302
QY 181 GTGAGCCCTCTCTATACCAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTGCCTCT 240
DB 303 GTGAGCCCTCTCTATACCAATGGCCCTCTCTCTCCAAATCCACAGGGAAGTGCCTCT 362
QY 241 GTATTGCAATGGTCTCGAGGACCTCTGTACCTGTGGCCCAATGGTGCCTGTGCCAC 300

DB 363 GTATTGCAATGGTCTCGAGGCTCTGTACCTGTGCCCAAAATGGTGCCTGTGTGCCAC 422
QY 301 CTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATTGATGCTTCTGTGAAGATCGTGAG 360
DB 423 CTGGTTTCAAGACCCCTACCGCTTCACTGGCACCATTGATGCTTCTGTGAAGATCGTGAG 482
QY 361 GCACGAGGACACAGGACCCCTCTGGAGCGCCCTCCCGCCACCCCTGGTGTGATGACTGTGCC 420
DB 483 GCACGAGGACACAGGACCCCTCTGGAGCGCCCTCCCGCCACCCCTGGTGTGATGACTGTGCC 542
QY 421 AGTACCGGCATCTACTTCACTGCTATGACCAACTGAAGGCCCTTCTGTGTGATGACTGTGAGC 480
DB 543 AGTACCGGCATCTACTTCACTGCTATGACCAACTGAAGGCCCTTCTGTGTGATGACTGTGAGC 602
QY 481 CCTGACCTCTGACCTCTAGCACCCTATGTCGCGCTGTGCGCTGGCCGCTGGGCAACCGT 540
DB 603 CCTGACCTCTGACCTCTAGCACCCTATGTCGCGCTGTGCGCTGGCCGCTGGGCAACCGT 662
QY 541 GACTGTGATCAGCCCTCTGGAGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGCTGA 600
DB 663 GACTGTGATCAGCCCTCTGGAGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGCTGA 722
QY 601 CCGGAGCTGGGTGCTGTGTTGAACTGAGTGGCTCAGGCTGGTGGCTGAGCTGCTGAGTGG 660
DB 723 CCGGAGCTGGGTGCTGTGTTGAACTGAGTGGCTCAGGCTGGTGGCTGAGCTGCTGAGTGG 782
QY 661 GCTGGGCTGGGCCCCCTGCTGCTGAGATGTGCTTCTC----- 702
DB 783 GCTGGGCTGGGCCCCCTGCTGCTGAGATGTGCTTCTCAGTGTATCCCCACCCCA 842
QY 703 AGCCCTCTACTGTTTCAACTATGAGTGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAA 762
DB 843 AGCCCTCTACTGTTTCAACTATGAGTGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAA 902
QY 763 GACACAGACTTCTGTGGGCTAGCTTGTGGCTGGTGGCTGTCATCTCAGGAGCGTGGCTGC 822
DB 903 GACACAGACTTCTGTGGGCTAGCTTGTGGCTGGTGGCTGTCATCTCAGGAGCGTGGCTGC 962
QY 823 AGTGCTGACTCTACCCCTTTGACGTGGTGAAGACCCCAAGCCAGGTGCTCTGGGAGCGAT 882
DB 963 AGTGCTGACTCTACCCCTTTGACGTGGTGAAGACCCCAAGCCAGGTGCTCTGGGAGCGAT 1022
QY 883 GGAGGCTGTGAGAGTGAACCCCTCTGATGTGGATCTCCACTGGCTGGTGTGCTGGGAGGAT 942
DB 1023 GGAGGCTGTGAGAGTGAACCCCTCTGATGTGGATCTCCACTGGCTGGTGTGCTGGGAGGAT 1082
QY 943 CCGGGCCGAGTGGGACCAAGAGGACTTTTTCAGAGCTTCTTCTCGATCATCAAGGC 1002
DB 1083 CCGGGCCGAGTGGGACCAAGAGGACTTTTTCAGAGCTTCTTCTCGATCATCAAGGC 1142
QY 1003 TGCCCCCTCTCTGTGCTATCATGATGACACCTATGATTTCCGCAAAAGCTTCTTCAGAG 1062
DB 1143 TGCCCCCTCTCTGTGCTATCATGATGACACCTATGATTTCCGCAAAAGCTTCTTCAGAG 1202
QY 1063 GCTGAACACGAGCCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGAGGAC 1114
DB 1203 GCTGAACACGAGCCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGAGGAC 1254
RESULT 3
AAC90452
ID AAC90452 standard; cDNA; 1560 BP.
XX AC AAC90452;
XX DT 12-MAR-2001 (first entry)
XX Human uncoupling protein cDNA #1.
DE Human; uncoupling protein; immunosuppressive; antiarthritic;
KW antineumatic; antiproliferative; cardiant; vasotropic;
KW cerebroprotective; neuroprotective; antibacterial; ophthalmological;
KW gastrointestinal; nephrotropic; gynaecological; vulnary; thrombolytic;

KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
 KW infertility; ss.
 XX Homo sapiens.
 OS
 PN WO200061614-A2.
 XX
 PD 19-OCT-2000.
 XX
 PF 06-APR-2000; 2000WO-US09534.
 XX
 PR 09-APR-1999; 99US-0128701.
 PR 08-JUL-1999; 99US-0142821.
 PR 18-AUG-1999; 99US-0149448.
 PR 12-NOV-1999; 99US-0164751.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;
 DR WPI; 2000-656322/63.
 DR P-PSDB; AAB50378.
 XX
 PT Uncoupling proteins and nucleic acid sequences encoding them, useful
 PT for detecting, preventing and treating proliferative, neurological,
 PT immune system, cardiovascular and gastrointestinal disorders -
 XX
 PS Claim 1; Page 303; 343pp; English.
 XX
 CC The present sequence is one of eighteen isolated nucleotide sequences
 CC encoding uncoupling proteins. The nucleotide sequences may be used for
 CC the detection of various disorders such as cancer, for chromosome
 CC identification, as chromosome markers and for numerous other diagnostic
 CC or research purposes. The uncoupling protein encoded by the nucleotide
 CC sequences may be used to treat disorders such as neural, immune,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders, wounds, infectious diseases,
 CC thrombosis, arthritis, and infertile y.
 XX
 SQ Sequence 1560 BP; 286 A; 492 C; 45 G; 323 T; 0 other;

Query Match 94.8%; Score 1056; DB 21; Length 1560;
 Best Local Similarity 97.8%; Pred. No. 8.5e-250;
 Matches 1090; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCCCTGCGGCGATCAGCCCTCCACCAATGGT 60
 DB 107 CTGAAGCTTCAAGATGGCTGACAGGACCCCTGCGGCGATCAGCCCTCCACCAATGGT 166
 QY 61 GGCCTCAGGACCGGGCTGTGGTACCTCTCTTATGACACCCCTGGAGTGGTGA 120
 DB 167 GGCCTCAGGACCGGGCTGTGGTACCTCTCTTATGACACCCCTGGAGTGGTGA 226
 QY 121 GGTTCGCTCAGTCTCAGGCGCTCCATGCGGAGGATGATGCTTCTCCAGACT 180
 DB 227 GGTTCGCTCAGTCTCAGGCGCTCCATGCGGAGGATGATGCTTCTCCAGACT 286
 QY 181 GTGAGGCTCTCTATACCAATATGCGCTCTCTCTCAATCCACAGGAGTGGCTCT 240
 DB 287 GTGAGGCTCTCTATACCAATATGCGCTCTCTCTCAATCCACAGGAGTGGCTCT 322
 QY 241 GTATTGCAATGGTGTCTGAGGCTCTGTACCTGTGCGGCGGCGGCTGGCGGCG 300
 DB 323 GTATTGCAATGGTGTCTGAGGCTCTGTACCTGTGCGGCGGCGGCTGGCGGCG 382
 QY 301 CTGGTTTCAAGACCTACCGCTCTACTGGCAGCATGATGCTTCTGTAAGATCGTGAG 360
 DB 383 CTGGTTTCAAGACCTACCGCTCTACTGGCAGCATGATGCTTCTGTAAGATCGTGAG 442
 QY 361 GCACGAGGCGACCGAGGACCTCTGAGGCGGCTCTCCCGCCACCTGGTGTGATGTGCC 420
 DB 443 GCACGAGGCGACCGAGGACCTCTGAGGCGGCTCTCCCGCCACCTGGTGTGATGTGCC 502

QY 421 AGCTACCCCATCTACTTCACTGCCTATGACCAACTGAAGCCCTTCTGTGTGTCGAGC 480
 DB 503 AGCTACCCCATCTACTTCACTGCCTATGACCAACTGAAGCCCTTCTGTGTGTCGAGC 562
 QY 481 CCTGACCTCTGACCTCTACGACCCCATGCTGGTGGCTGGCGCGCTGGCGCGCTGGCGCGCT 540
 DB 563 CTTGACCTCTGACCTCTACGACCCCATGCTGGTGGCTGGCGCGCTGGCGCGCTGGCGCGCT 622
 QY 541 GACTGTGATCAGCCCTCTGAGCTTATGCGGACAAAGCTGCAAGCTTCAGCTTGTGCTGA 600
 DB 623 GACTGTGATCAGCCCTCTGAGCTTATGCGGACAAAGCTGCAAGCTTCAGCTTGTGCTGA 682
 QY 601 CCGGAGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 DB 683 CCGGAGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 742
 QY 661 GCTGGGCTGGGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 743 GCTGGGCTGGGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
 QY 721 CTATGAGCTGGTGAAGAGCTGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGGG 780
 DB 803 CTATGAGCTGGTGAAGAGCTGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGGG 862
 QY 781 CATGAGCTTGTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 863 CATGAGCTTGTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922
 QY 841 TGACCTGTGAAGAGCTGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGGG 900
 DB 923 TGACCTGTGAAGAGCTGCTCAATGGGCTCAGGCGGAGGACCAAGCTTCTGTGGG 982
 QY 901 CCCCCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 DB 983 CCCCCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1042
 QY 961 CAGGAGCTTCTTCCAGGCTTCTTCTTCCGATCATCAAGGCTGCGGCGGCTGCTGCTGCTGCTGCT 1020
 DB 1043 CAGGAGCTTCTTCCAGGCTTCTTCTTCCGATCATCAAGGCTGCGGCGGCTGCTGCTGCTGCTGCT 1102
 QY 1021 CATGATCAGCACCCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1080
 DB 1103 CATGATCAGCACCCTATGAGTTCGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1162
 QY 1081 TCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
 DB 1163 TCTGGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1196
 RESULT 4
 AAI71020
 ID AAI71020 standard; cDNA; 1546 BP.
 XX
 AC AAI71020;
 VT 18-MAR-2002 (first entry)
 XX
 DE Human mitochondrial carrier protein CGI-69 cDNA.
 KW CGI-69; mitochondrial carrier protein; human; metabolic disease;
 KW obesity; cachexia; tumour; cancer; infection; immunomodulator;
 KW antitumour; virucide; antibacterial; anorectic; antidiabetic;
 KW brown adipose tissue; diagnosis; gene therapy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 118..1173
 FT /*tag= a
 XX
 FN WO200198355-A2.
 XX
 PD 27-DEC-2001.

xx PF 22-JUN-2001; 2001WO-US20117.
xx PR 22-JUN-2000; 2000US-213307P.
xx (GETH) GENENTECH INC.
xx PA (CURA-) CURAGEN CORP.
xx PL Lewin D, Adams SH, Yu XX;
xx DR WPI; 2002-114569/15.
xx CGI-69 polypeptides and polynucleotides useful for treating metabolic
xx disorders including cachexia, obesity, diabetes and cancers -
xx Disclosure; Page 11-12; 107pp; English.
xx The present sequence is that of cDNA encoding human CGI-69 (see
xx AM50570). Various CGI-69 clones were isolated from human liver
xx upon PCR amplification and cloning. Some diverged from the
xx previously known CGI-69 sequence in that they encoded an 8-amino
xx acid insert preceded by a W6L change, and were termed the 'long
xx version' isoform or CGI-69L (see AAI71019). The invention relates
xx to a novel characterization of CGI-69 as a mitochondrial carrier
xx protein, the discovery of CGI-69L, and the discovery of the
xx alteration of the mitochondrial membrane potential through
xx overexpression of carboxyl-FLAG-tagged CGI-69. CGI-69 (including
xx splice variant) nucleic acids and proteins are useful for
xx diagnosing and treating metabolic diseases in humans, including
xx obesity, cachexia and diabetes. Evidence for an important function
xx for CGI-69 in modifying mitochondrial membrane potential in brown
xx adipose tissue (BAT) is presented. The mouse orthologue of CGI-69
xx is up-regulated in cold-treated BAT. CGI-69 may be involved in
xx cellular thermogenic uncoupling and, therefore, may be used to
xx diagnose and treat specific perturbations in metabolic pathways.
xx Altering the expression of CGI-69 through gene therapy provides a
xx means of treating metabolic diseases, such as obesity or cachexia,
xx or of increasing or decreasing body weight. Decreasing the
xx activity of CGI-69 can be used to treat cachexia, tumours, cancers,
xx viral infections and bacterial infections, while increasing its
xx activity can be used to treat obesity, tumours, cancers, viral
xx infections and bacterial infections (all claimed). Methods are
xx also claimed for determining whether a compound up-regulates or
xx down-regulates expression of a CGI-69 gene, and of screening for
xx a mutation in the CGI-69 gene.
xx Sequence 1546 BP; 272 A; 489 C; 459 G; 326 T; 0 other;
ry Match 94.6%; Score 1054.4; DB 24; Length 1546;
Jt Local Similarity 97.8%; Pred. No. 2.1e-249;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 1 CTGAAGCTTCAGATGGTGTACAGGACCTTGGGGCATCAGCCCTCCAGCAATGCT 60
Db 105 CTGAAGCTTCAGATGGTGTACAGGACCTTGGGGCATCAGCCCTCCAGCAATGCT 164
QY 61 GGCTCAGGACCGGGGTGTGGTACCTCTCTTCATGACACCCCTGGAGGTGTA 120
Db 165 GGCTCAGGACCGGGGTGTGGTACCTCTCTTCATGACACCCCTGGAGGTGTA 224
QY 121 GGTTCGCTGCAGTCTCAGCGGCCCTCCATGGCCAGCGAGTGTGCTTCTCCAGACT 180
Db 225 GGTTCGCTGCAGTCTCAGCGGCCCTCCATGGCCAGCGAGTGTGCTTCTCCAGACT 284
QY 181 GTGAGCCTCTCTATACCAATTGCCCTCTCTCTCCATCCAGGGAAGTGCCTCT 240
Db 285 GTGAGCCTCTCTATACCAATTGCCCTCTCTCTCCATCCAGGGAAGTGCCTCT 320
QY 241 GTATTGCAATGTGTCTCGAGGCTGTGTACCTGTGCCCAATGTGTGCTGTGCGAC 300
Db 321 GTATTGCAATGTGTCTCGAGGCTGTGTACCTGTGCCCAATGTGTGCTGTGCGAC 380
QY 301 CTGGTTTCAAGACCCCTACCCGCTTCACTGGCACCATGATGCTTCTGTGAAGATCGTGAG 360

Db 381 CTGGTTTCAAGACCCCTACCCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCGTGAG 440
QY 361 GCACGAGGGGACACGAGACCTCTGAGCGGCTCCCGCCACCCCTGTGTGATGACTGTGCC 420
Db 441 GCACGAGGGGACACGAGACCTCTGAGCGGCTCCCGCCACCCCTGTGTGATGACTGTGCC 500
QY 421 AGCTACCGGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGCGAGC 480
Db 501 AGCTACCGGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTGCGAGC 560
QY 481 CCTGACCTCTGACCTCTACGACCCCATGTTGGCTGGCGGCTTGGCCGCTTGGSCACCGT 540
Db 561 CCTGACCTCTGACCTCTACGACCCCATGTTGGCTGGCGGCTTGGCCGCTTGGSCACCGT 620
QY 541 GACTGTGATCAGCCCTCTGAGCTTATCGGACAAAGCTGAGGCTCAGCATGTGTGCTA 600
Db 621 GACTGTGATCAGCCCTCTGAGCTTATCGGACAAAGCTGAGGCTCAGCATGTGTGCTA 680
QY 601 CCGGAGCTGGGTGGCTGTGTTCGAACTGCAGTGGCTCAGGCTGGCTGGCGCTCAGCTGTG 660
Db 681 CCGGAGCTGGGTGGCTGTGTTCGAACTGCAGTGGCTCAGGCTGGCTGGCGCTCAGCTGTG 740
QY 661 GCTGGGCTGGGCGCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAA 720
Db 741 GCTGGGCTGGGCGCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAA 800
QY 721 CTATGAGCTGGTGAAGACCTGGCTCAATGGGCTCAGGCCGGAAGACAGACTTCTGTGG 780
Db 801 CTATGAGCTGGTGAAGACCTGGCTCAATGGGCTCAGGCCGGAAGACAGACTTCTGTGG 860
QY 781 CATGAGCTTGTGGCTGGCTGCTCATCTCAGGACGGTGGCTGCAGTGGCTGACTCTACCCCT 840
Db 861 CATGAGCTTGTGGCTGGCTGCTCATCTCAGGACGGTGGCTGCAGTGGCTGACTCTACCCCT 920
QY 841 TGACGTGTGAAGACCCCAAGCCAGGCTGGCTCTGGAGCGATGAGGCTGTGAGAGTGA 900
Db 921 TGACGTGTGAAGACCCCAAGCCAGGCTGGCTCTGGAGCGATGAGGCTGTGAGAGTGA 980
QY 901 CCCCTGTCATGTGACTCAGCTGGCTGCTGGAGGATCCGGGCCGAGTCGGGCAC 960
Db 981 CCCCTGTCATGTGACTCAGCTGGCTGCTGGAGGATCCGGGCCGAGTCGGGCAC 1040
QY 961 CAAGGACTCTTTCAGAGGCTTCTCTTCGAGTATCAAGGCTGCCCTCTCTGTGCTAT 1020
Db 1041 CAAGGACTCTTTCAGAGGCTTCTCTTCGAGTATCAAGGCTGCCCTCTCTGTGCTAT 1100
QY 1021 CATGATCAGCACTATGAGTTGGCAAAAGCTTCTTTCAGAGGCTGAACACGAGCCGGCT 1080
Db 1101 CATGATCAGCACTATGAGTTGGCAAAAGCTTCTTTCAGAGGCTGAACACGAGCCGGCT 1160
QY 1081 TCTGGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 1114
Db 1161 TCTGGGCGGCTGAAGGGGCAAGGAGGCAAGGAC 1194
RESULT 5
ABI99647
ID ABI99647 standard; cDNA; 704 BP.
XX AC ABI99647;
XX DT 07-MAR-2002 (first entry)
XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:675.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX OS Mus musculus.
XX PN W0200188188-A2.
XX

PD 22-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-JP04192.
 XX
 PR 18-MAY-2000; 2000JP-0145977.
 XX
 PA (UNIV-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX WPI; 2002-034733/04.
 XX
 DR
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes -
 XX
 PS Claim 2; Page 1679-1680; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI9202 to ABI9912, encoding
 CC the protein sequences in AB57020 to AB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI9913 and ABI9914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.
 XX
 SQ Sequence 704 BP; 132 A; 204 C; 209 G; 159 T; 0 other;
 Query Match 42.4%; Score 472; DB 24; Length 704;
 Best Local Similarity 83.2%; Pred. No. 2.3e-106;
 Matches 561; Conservative 0; Mismatches 110; Indels 3; Caps 2;
 QY 432 TCTACTTCAGCTGCTATGACCACTGAAGGCTTCCTGTGTGTGCTGAGCCCTGACCTCTG 491
 DB 3 TCGATCCATGGCTTACGACCACTCAAGGCTTCCTGTGTGTGCTGACCTTGACCTCTG 62
 QY 492 ACCTTACGACCATGCTGGTGGCGCTGGCCGCTGGCCGCTGGCCGCTGGCTGATGATCA 551
 DB 63 ACCTTACGACCATGCTGGTGGCGCTGGCCGCTGGCCGCTGGCTGATGATGATCA 120
 QY 552 GCGGCTGGAGCTTATGCGGACCAAGCTGCAGGCTCAGCATGTGCTACCGGAGCTGG 611
 DB 121 GCGGCTGGAGCTTATGCGGACCAAGCTGCAGGCTCAGCATGTGCTACCGGAGCTGG 180
 QY 612 GTGCTGTGTTTGAATCTGAGTGGCTGAGGCTGGCTGCTGCTGCTGCTGCTGCTGG 671
 DB 181 CTTCTCTGTTCAAGCTGCGGTGACTCAGGCTGGCTGGCGCTCTCTGTGCTGGGCTGG 240
 QY 672 GCGGCTGGCTTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 731
 DB 241 GTCCACAGCTCTTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 QY 732 TGAAGAGTGGCTCAATGGCTCAGCGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
 DB 301 TGAAGAGTGGCTCAGCGAGCTGAGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 QY 792 TGGCTGTGTTGCTCAGGACGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 851
 DB 361 TGGCGGCTGG-ATCTAGGAATGGTGGCTGGCCACCTTACCTTACCTTACCTTACCTTAC 419
 QY 852 AGACCCAAACCGGCTGCTTGGAGGCTGAGGCTGCTGAGTGAACCCCTTGCATG 911
 DB 420 AGACACAGGACACATGCTCACTGGGAGGCTGAGGCTGCTGAGAGTGAAGCCGCGCAGAG 479
 QY 912 TGGACTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971

Db 480 TGGACTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 QY 972 TTGAGAGCTTCCTTCTCGGATCATCAAGGCTGCCCCCTCTGCTGCTGCTGCTGCTGCTGCT 1031
 Db 540 TTGAGAGCTTCCTTCTCGGATCATCAAGGCTGCCCCCTCTGCTGCTGCTGCTGCTGCTGCT 599
 QY 1032 CTTATGATGCTGCGCAAAAGCTTCTTCCAGAGGCTGAACACGAGGCTTCTGCGGCGCT 1091
 Db 600 CTTACGAGTTTGGCAAAAGCTTCTTCCAGAGGCTCAACACGAGGAGGAGCTTCTTGGCGGCT 659
 QY 1092 GAAAGGGGCAAGCA 1105
 Db 660 TGAAGTGGGAGCA 673
 RESULT 6
 ID ABK34735 standard; cDNA; 2170 BP.
 XX
 AC ABK34735;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA for novel secreted protein, SEQ ID 504.
 XX
 KW Human; sr; gene; secreted protein; immune deficiency; viral infection;
 KW bacterial infection; fungal infection; autoimmune disorder; burn;
 KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
 KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
 KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
 KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
 KW lymphoid cell deficiency.
 XX
 OS Homo sapiens.
 XX
 PN WO200177290-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US10295.
 XX
 PR 06-APR-2000; 2000US-194941P.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 Gulukota K, Graham JR;
 XX WPI; 2002-179323/23.
 XX
 PT Six hundred and twenty five polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for
 PT treating immune deficiencies and disorders such as autoimmune disorders
 PT -
 XX
 PS Claim 1; Page 269-270; 339pp; English.
 XX
 CC The invention relates to 625 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins, their complements and sequences that hybridise to them.
 CC Also included are a vector comprising the polynucleotide, a host cell
 CC transformed with the vector, the proteins encoded by the
 CC polynucleotides, antibodies that bind to the proteins and identification
 CC of modulators of the proteins or the expression of the polynucleotide.
 CC The polynucleotides can be used as probes for the identification
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 CC and proteins can also be used as nutritional supplements. The protein
 CC is useful in the treatment of various immune deficiencies and disorders
 CC such as viral infections, bacterial infections, fungal infections,
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 CC autoimmune thyroiditis and diabetes) and allergic reactions and

QY 1062 GGCTGAA 1068
 Db 2731 AACAAAA 2725

RESULT 9

ABL81508
 ID ABL81508 standard; cDNA; 453 BP.

XX
 AC ABL81508;

XX 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:4486.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US17756.

XX 26-MAY-2000; 2000US-207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide -

XX Claim 1; SEQ ID 4486; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers and immunostimulants; and a polypeptide (III), of a ovarian tumour polypeptide encoded by a polynucleotide (III) having a cDNA sequence (SI) from the 10912 nucleotide sequences as given in ABL77023 to ABL87934, (III) encoding (II) having a sequence (S2), a T cell population of (II), or antigen presenting cells that express (II). (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to (SI) can be used for detecting ovarian cancer in a patient's biological sample preferably serum or ovarian tissue. The method comprises contacting a biological sample from a patient with (IV), detecting the amount of polynucleotide hybridising to (IV) and comparing the amount to a predetermined cutoff value and thereby detecting ovarian cancer in the patient, where the amount of polynucleotide hybridising to (IV) is detected preferably by polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is useful for stimulating and/or expanding T cells specific for an ovarian tumour protein comprising contacting T cells with (III) or (II). (III) is useful in design and preparation of ribozyme molecules for inhibiting expression of the tumour polypeptides and proteins in tumour cells; and to isolate a full length gene from a suitable library e.g., a tumour cDNA library using well known techniques.

XX Sequence 453 BP; 74 A; 154 C; 128 G; 97 T; 0 other;

Query Match 16.7%; Score 185.8; DB 24; Length 453;
 Best Local Similarity 92.0%; Pred. No. 4.5e-36;
 Matches 207; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 503 CCATGTCGTCGGCGCGCTGGCCCGCTGGGACCGTGTGATCAGCCCGCTGGAG 562

Db 230 CCATGTCACCCCAATCCCAACACAGTGGGACCGTGTGATCAGCCCGCTGGAG 289

QY 563 CTTATCGGACAAAGCTGAGGCTCAGCATGTGTCACCGGAGCTGGGTCCTGTGTT 622

Db 290 CTTATCGGACAAAGCTGAGGCTCAGCATGTGTCGATCGGAGCTGGTGCCTGTGTT 349
 QY 623 CGAACTGCAGTGGCTCAGGTCGGCTGGCGCTCACTGGCTGGGTCGGGCCCCACTGCC 682
 Db 350 CGAACTGCAGTGGCTCAGGTCGGCTGGCGCTCACTGGCTGGGTCGGGCCCCACTGCC 408
 QY 683 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 727
 Db 409 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 453

RESULT 10

ABK62422/c

ID ABK62422 standard; cDNA; 512 BP.

XX
 AC ABK62422;

XX 18-JUN-2002 (first entry)

XX Rat sequence differentially expressed in response to a hepatotoxin #329.

XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;

XX differential expression; centrilobular necrosis; steatosis.

XX Rattus norvegicus.

XX WO200210453-A2.

XX 07-FEB-2002.

XX 30-JUL-2001; 2001WO-US23872.

XX 31-JUL-2000; 2000US-222040P.

XX 02-NOV-2000; 2000US-244880P.

XX 11-MAY-2001; 2001US-290029P.

XX 15-MAY-2001; 2001US-290645P.

XX 22-MAY-2001; 2001US-292336P.

XX 06-JUN-2001; 2001US-295798P.

XX 13-JUN-2001; 2001US-297457P.

XX 19-JUN-2001; 2001US-298884P.

XX 09-JUL-2001; 2001US-303459P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;

XX WPI; 2002-241625/29.

XX Claim 1; Seq ID No 329; 239pp; English.

XX The invention relates to methods for predicting toxic effects of compounds or the progression of these toxic effects by determining the global changes in gene expression in tissues or cells exposed to the toxin and comparing these to gene expression in unexposed tissues or cells. Also included are methods of predicting at least one toxic effect of a compound or progression of a toxic effect, preferably the hepatotoxicity of a compound, comprising detecting the level of expression in a tissue or cell sample exposed to the compound of two or more genes listed in the specification, where differential expression of the genes is indicative of at least one toxic effect or progression. The method can also be used to identify an agent which modulates the toxic response and predict cellular pathways that a compound modulates in a cell. The methods utilise a set of at least two probes (on a solid support in kit form), where each of the probes comprises a sequence that specifically hybridises to a gene listed in the specification, a computer system comprising a database containing information identifying the expression level in a tissue or cell sample exposed to a hepatotoxin of a

CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information,
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.
 XX
 SQ Sequence 512 BP; 119 A; 129 C; 147 G; 116 T; 1 other;

Query Match 15.3%; Score 170.2; DB 24; Length 512;
 Best Local Similarity 81.7%; Pred. No. 3.1e-32;
 Matches 196; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 857 CAACGCCAGTCTGCTGGAGCATGGAGCTGTGAGAGTGAACCCCTGCATGTGGAC 916

Db 511 CAGCGCAAAATGTCACTGGAGCAGTGGAGGCTATGAGAGTGAAGCCGCAAGAGTTGAC 452

QY 917 TCCACCTGGTCTGCTGGAGGATCGGGCGAGTCCGGCACCACCAAGGACTCTTTGCA 976

Db 451 TCCACCTGGTCTGCTGGAGGATCGGGCGAGTCCGGCACCACCAAGGACTCTTTGCA 392

QY 977 GGCTTCCTTCCTGGATCATCAAGGCTGCCCTCTCTGTGCCATCATGATCAGCACCTAT 1036

Db 391 GGTTCCTTCCTGGATCATCAAGGCTGCCCTCTCTGTGCCATCATGATCAGCACCTAT 332

QY 1037 GAGTTGGGAAAAGCTTCTTCAGAGGCTGAACAGAGCAGCGCTTCTGGCGCTGAAG 1096

Db 331 GAGTTGGGAAAAGCTTCTTCACAGGCTCAACAGGAGCAGCCTCTGGCGCTGAAG 272

RESULT 11

AAA06606
 ID AAA06606 standard; cDNA; 537 BP.

XX
 AC AAA06606;

XX
 DT 13-JUN-2000 (first entry)

XX
 DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:387.

XX
 DE Human; prostate cancer; diagnosis; tumour; gene therapy; detection;

XX
 KW immunogenic; cytostatic; vaccine; ss.

XX
 OS Homo sapiens.

XX
 PN WO200004149-A2.

XX
 PD 27-JAN-2000.

XX
 PF 14-JUL-1999; 99WO-US15838.

XX
 PR 14-JUL-1998; 98US-0115453.

XX
 PR 14-JUL-1998; 98US-0116134.

XX
 PR 23-SEP-1998; 98US-0159812.

XX
 PR 23-SEP-1998; 98US-0159822.

XX
 PR 15-JAN-1999; 99US-0232149.

XX
 PR 15-JAN-1999; 99US-0232880.

XX
 PR 09-APR-1999; 99US-0288946.

XX
 PA (CORI-) CORIXA CORP.

XX
 PI Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;

XX
 DR WPI; 2000-171268/15.

XX
 PT New polypeptide useful for treating and diagnosing prostate cancer

PT comprises an immunogenic portion of prostate tumor protein -
 XX
 PS Claim 50; Page 235-236; 263pp; English.

XX
 CC The present invention describes isolated polypeptides, comprising an
 CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
 CC and polynucleotides encoding them have cytostatic activity and can be
 CC used in vaccines and in gene therapy. The polypeptides and
 CC polynucleotides encoding them, antigen presenting cells which express
 CC the polypeptides, antibodies against the polypeptides and vaccines
 CC comprising them can be used for inhibiting the development of prostate
 CC cancer in a patient. The polypeptides can be used to generate antibodies
 CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
 CC the polynucleotides encoding the polypeptides can be used as a probe or
 CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
 CC AA82000 to AA82020 represent sequences used in the exemplification of
 CC the present invention.

XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;

Query Match 15.3%; Score 170; DB 21; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCGAGTCGGCACCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAGGCTG 1004

Db 1 GGGCGAGTCGGCACCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAGGCTG 60

QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 1064

Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGATTCGGCAAAAGCTTCTTCCAGAGGC 120

QY 1065 TGAACCAAGACCGGCTTCTGGCGCTGAAAGGGGCAAGGCAAGGCAAGGAC 1114

Db 121 TGAACCAAGACCGGCTTCTGGCGCTGAAAGGGGCAAGGCAAGGCAAGGAC 170

RESULT 12

AA563815
 ID AA563815 standard; cDNA; 537 BP.

XX
 AC AA563815;

XX
 DT 29-JAN-2002 (first entry)

XX
 DE Human prostate cDNA sequence #357.

XX
 DE Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.

XX
 OS Homo sapiens.

XX
 PN WO200173032-A2.

XX
 PD 04-OCT-2001.

XX
 PF 27-MAR-2001; 2001WO-US09919.

XX
 PR 27-MAR-2000; 2000US-0536857.

XX
 PR 09-MAY-2000; 2000US-0568100.

XX
 PR 12-MAY-2000; 2000US-0570737.

XX
 PR 13-JUN-2000; 2000US-0593793.

XX
 PR 27-JUN-2000; 2000US-0605783.

XX
 PR 10-AUG-2000; 2000US-0636215.

XX
 PR 29-AUG-2000; 2000US-0651236.

XX
 PR 06-SEP-2000; 2000US-0657279.

XX
 PR 02-OCT-2000; 2000US-0679426.

XX
 PR 10-OCT-2000; 2000US-0685166.

XX
 PA (CORI-) CORIXA CORP.

XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2001-639232/73.
 XX
 PT New human prostate-specific polypeptides and polynucleotides useful for
 PT the diagnosis and treatment of cancer, especially prostate cancer -
 XX
 PS Claim 1; Page 361; 579pp; English.
 XX
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells specific for a tumour protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polynucleotide of the invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 15.3%; Score 170; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 945 GGGCCGAGTCGGCACCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 1004
 Db 1 GGGCCGAGTCGGCACCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 60
 QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGGGCAAAAGCTTCTTCCAGAGGC 1064
 Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGGGCAAAAGCTTCTTCCAGAGGC 120
 QY 1065 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCGAAGGAGCAAGGAC 1114
 Db 121 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCGAAGGAGCAAGGAC 170
 RESULT 13
 AAH93722
 ID AAH93722 standard; cDNA; 537 BP.
 AC AAH93722;
 DT 04-OCT-2001 (first entry)
 DE Human prostate-specific cDNA sequence CGI-69.
 KW Human; prostate-specific cDNA sequence CGI-69.
 KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
 KW cytostatic; gene therapy; metastasis; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200151633-A2.
 PD 19-JUL-2001.
 PF 16-JAN-2001; 2001WO-US01574.
 XX 14-JAN-2000; 2000US-0483672.
 PR (CORI-) CORIXA CORP.
 PA Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
 PI Wang A, Meagher MJ;
 XX WPI; 2001-425873/45.
 DR
 XX New polynucleotide encoding a prostate-specific protein, for

PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines -
 XX
 PS Claim 1; Page 359-360; 543pp; English.
 XX
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
 CC AAH01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 15.3%; Score 170; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 945 GGGCCGAGTCGGCACCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 1004
 Db 1 GGGCCGAGTCGGCACCACCAAGGACTCTTTCAGGCTTCCTTCGATCATCAAGGCTG 60
 QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGGGCAAAAGCTTCTTCCAGAGGC 1064
 Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGGGCAAAAGCTTCTTCCAGAGGC 120
 QY 1065 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCGAAGGAGCAAGGAC 1114
 Db 121 TGAACCAAGGACCGCTTCGGCGGCTGAAAGGGCGAAGGAGCAAGGAC 170
 RESULT 14
 AAH85036
 ID AAH85036 standard; cDNA; 537 BP.
 AC AAH85036;
 DT 25-SEP-2001 (first entry)
 DE Human prostate-specific cDNA sequence CGI-69.
 KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
 KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
 KW prostate specific antigen; PSA; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200134802-A2.
 PD 17-MAY-2001.
 PF 09-NOV-2000; 2000WO-US00904.
 XX 12-NOV-1999; 99US-0439313.
 PR 18-NOV-1999; 99US-0443686.
 XX (CORI-) CORIXA CORP.
 PA Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, wang A;
 DR WPI; 2001-308785/32.
 XX
 PT Isolated polypeptide comprising at least an immunogenic portion of a
 PT prostate-specific protein, useful in the diagnosis and therapy of

PT prostate cancer -
 XX Claim 31; Page 259; 325pp; English.
 XX
 CC The present invention describes an isolated polypeptide (P1) comprising
 CC at least an immunogenic portion of a prostate-specific protein, or its
 CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
 CC (N1) have cytostatic activity and can be used in vaccine production.
 CC The polypeptides, nucleic acids and antibodies from the present
 CC invention are useful in the diagnosis and therapy of prostate cancer.
 CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
 CC region. Prostate specific antigen (PSA) P501S was located on
 CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
 CC polynucleotide and polypeptide sequences used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 15.3%; Score 170; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 945 GGCCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
 DB 1 GGCCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
 QY 1005 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 1064
 DB 61 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 120
 QY 1065 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGAC 1114
 DB 121 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGAC 170
 RESULT 15
 AAH02787
 ID AAH02787 standard; cDNA; 537 BP.
 AC AAH02787;
 DT 14-JUN-2001 (first entry)
 XX
 DE Prostate tumour antigen cDNA sequence for CGI-69.
 KW Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
 KW prostate cancer; immunogenic; cytostatic; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200125272-A2.
 PD 12-APR-2001.
 XX
 PF 04--CT-2000; 2000WO-US27464.
 XX
 PR 04--O-T-1999; 99US-0157455.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Skelky YAW, Reed SG, Cheever MA;
 DR WPI; 2001-245062/25.
 XX
 PT Prostate specific protein and its encoding polynucleotide, useful for
 PT the treatment and diagnosis of prostate cancer -
 XX
 PS Claim 50; Page 244; 276pp; English.
 XX
 CC The present invention describes an isolated polypeptide (I) comprising
 CC at least an immunogenic portion of a prostate tumour antigen protein or
 CC its variant. (I) have cytostatic activity and can be used in vaccine

CC production. (I), prostate tumour antigen polynucleotides, an antigen
 CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
 CC pharmaceutical composition containing (I) are useful for inhibiting the
 CC development of cancer in a patient. Antibodies specific for prostate
 CC specific proteins and oligonucleotides that hybridise to a
 CC polynucleotide that encodes a prostate specific protein are useful
 CC for detecting the presence or absence of a cancer or monitoring the
 CC progression the progression of a cancer, especially prostate cancer.
 CC AAH02422 to AAH2872, AAH74798 to AAH74821 and AAH74830 are sequences
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 537 BP; 116 A; 167 C; 138 G; 116 T; 0 other;
 Query Match 15.3%; Score 170; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 3.6e-32;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 945 GGCCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
 DB 1 GGCCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
 QY 1005 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 1064
 DB 61 CCCCTCTCTGTGCCATCATGATCAGCACCTATGAGTTGGCAAAAGCTTCTTCAGAGGC 120
 QY 1065 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGAC 1114
 DB 121 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGAGGCAAGGAC 170

Search completed: January 22, 2003, 19:03:32
 Job time : 280.466 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 16:32:11 ; Search time 45.6489 seconds
(without alignments)
7484.038 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagcttcaagatgctg.....aggggcaaggaggaaggac 1114

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 1533381 residues

T. number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1054.4	94.6	1643	2	US-08-933-750C-68
2	1054.4	94.6	1643	3	US-09-234-613-68
3	170	15.3	537	4	US-09-605-785-387
4	170	15.3	537	4	US-09-439-313-387
5	170	15.3	537	4	US-09-352-616A-387
6	61	5.5	997	3	US-09-188-930-23
7	61	5.5	1816	3	US-09-188-930-262
8	48.8	4.4	1311	4	US-09-068-140A-9
9	48.8	4.4	5150	4	US-09-068-140A-14
10	47.6	4.3	7218	1	US-08-232-463-14
11	42.2	3.8	1594	2	US-08-933-750C-61
12	42.2	3.8	1594	3	US-09-234-613-61
13	40.8	3.7	1259	3	US-08-961-871-11
14	39.8	3.6	1192	4	US-09-142-565-1
15	39.6	3.6	897	4	US-09-434-288-6
16	39.2	3.5	1177	3	US-08-961-871-9
17	39	3.5	2634	1	US-08-196-218-31
18	39	3.5	2634	1	US-08-681-953-31
19	36.8	3.3	8460	1	US-08-469-005A-9
20	35.6	3.2	1535	1	US-08-910-973-10
21	35.6	3.2	1535	4	US-09-499-227-10
22	35.6	3.2	1550	3	US-09-234-332-3
23	35.6	3.2	1722	4	US-09-385-028-15
24	35.6	3.2	11604	4	US-09-385-028-13
25	35.6	3.2	15079	4	US-09-385-028-1
26	35.4	3.2	606	2	US-08-432-871C-48
27	35.4	3.2	606	4	US-09-270-956-48

28	35.4	3.2	2681	4	US-08-928-213B-7	Sequence 7, Appli
29	35.2	3.2	50341	1	US-08-247-901C-1	Sequence 1, Appli
30	35.2	3.2	50341	2	US-09-075-904-1	Sequence 1, Appli
31	35.2	3.2	52297	4	US-09-426-436-1	Sequence 1, Appli
32	35.2	3.2	52297	4	US-08-705-557-1	Sequence 1, Appli
33	35.2	3.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
34	34.8	3.1	1949	2	US-08-937-466-3	Sequence 3, Appli
35	34.8	3.1	1949	2	US-09-172-528-3	Sequence 3, Appli
36	34.8	3.1	1949	3	US-09-318-199-3	Sequence 3, Appli
37	34.8	3.1	1949	4	US-09-503-579-3	Sequence 3, Appli
38	34.6	3.1	1421	4	US-08-943-731-207	Sequence 207, App
39	34.6	3.1	20084	4	US-08-943-731-5	Sequence 5, Appli
40	34.2	3.1	34303	2	US-08-735-609-4	Sequence 4, Appli
41	34.2	3.1	34303	2	US-08-735-609-4	Sequence 4, Appli
42	34.2	3.1	34303	3	US-09-315-372-4	Sequence 4, Appli
43	34.2	3.1	34303	3	US-09-244-752-4	Sequence 4, Appli
44	34.2	3.1	34303	3	US-09-245-497-4	Sequence 4, Appli
45	34.2	3.1	34303	4	US-09-562-919-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-933-750C-68
; Sequence 68, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNCOAT01

CLONE: 724157
US-08-933-750C-68

Query Match 94.6%; Score 1054.4; DB 2; Length 1643;
Best Local Similarity 97.8%; Pred. No. 3.9e-264;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGATCAGCCCTCCAGCAAAATGGT 60
DB 192 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGATCAGCCCTCCAGCAAAATGGT 251

QY 61 GGCCTCAGGACCGGGCTGTGTACTCTCTTCATGACACCCCTGGAGCTGTGA 120
DB 252 GGCCTCAGGACCGGGCTGTGTACTCTCTTCATGACACCCCTGGAGCTGTGA 311

QY 121 GGTTCGGCTGCGAGTCTCAGCGCCCTCCATGGCCAGCGAGCTGATGCTTCTCCAGACT 180
DB 312 GGTTCGGCTGCGAGTCTCAGCGCCCTCCATGGCCAGCGAGCTGATGCTTCTCCAGACT 371

QY 191 GTGAGGCTCTCTATACCAAAATGGCCCTCTCTCTCCAAATCCAGGGAAGTGCCTCT 240
DB 372 GTGAGGCTCTCTATACCAAAATGGCCCTCTCTCTCCAAATCCAGGGAAGTGCCTCT 407

QY 241 GTATTGAATGGTGTCTGGAGCTCTGTACTGTGCCAAATGGTGGCCGCTGTGCCAC 300
DB 408 GTATTGAATGGTGTCTGGAGCTCTGTACTGTGCCAAATGGTGGCCGCTGTGCCAC 467

QY 301 CTGTTTCAAGACCTTACCGCTTCACTGACCATGATGCTTCTGTGAAGATCGTGAG 360
DB 468 CTGTTTCAAGACCTTACCGCTTCACTGACCATGATGCTTCTGTGAAGATCGTGAG 527

QY 361 GCAGGAGCAGGACGACCTCTGAGGCGCTCCCGCCACCTGGTGCATGACTGTGC 420
DB 528 GCAGGAGCAGGACGACCTCTGAGGCGCTCCCGCCACCTGGTGCATGACTGTGC 587

QY 421 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGCTGAGC 480
DB 588 AGCTACCGCATCTACTTCACTGCTATGACCAACTGAAGGCTTCTGTGTGTGCTGAGC 647

QY 481 CTTGACCTTGACCTTACGACCATGATGGTGTGCGCTGGCCCGCTGGGCAACCGT 540
DB 648 CTTGACCTTGACCTTACGACCATGATGGTGTGCGCTGGCCCGCTGGGCAACCGT 707

QY 541 GACTGTGATCAGCCCTTGAGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGCTGA 600
DB 708 GACTGTGATCAGCCCTTGAGCTTATGCGGACAAAGCTGAGGCTCAGCATGTGCTGA 767

QY 601 CCGGAGCTGGGTGCTGTGTTCGAACTGACAGTGGCTCAGGGTGGCTGGCCCTCACTGTG 660
DB 768 CCGGAGCTGGGTGCTGTGTTCGAACTGACAGTGGCTCAGGGTGGCTGGCCCTCACTGTG 827

QY 661 GCTGGGTGGGGCCCACTGCGCTTTCAGATGTGCGCTTCTCAGCCCTGTACTGTTCAA 720
DB 828 GCTGGGTGGGGCCCACTGCGCTTTCAGATGTGCGCTTCTCAGCCCTGTACTGTTCAA 887

QY 721 CTATGAGCTGGTGAAGAGCTGCTCAATGGGCTCAGGCGGAGGACAGACTCTGTGGG 780
DB 888 CTATGAGCTGGTGAAGAGCTGCTCAATGGGCTCAGGCGGAGGACAGACTCTGTGGG 947

QY 781 CATGAGCTTGTGGCTGTGGCTCTCAGGAGCGGTGGCTGCGAGTGTGACTTACCTT 840
DB 948 CATGAGCTTGTGGCTGTGGCTCTCAGGAGCGGTGGCTGCGAGTGTGACTTACCTT 1007

QY 841 TGACCTGGTAAAGACCCCAAGCTGGTCTGTGGAGCGATGGAGGTGTGAGGTGA 900
DB 1008 TGACCTGGTAAAGACCCCAAGCTGGTCTGTGGAGCGATGGAGGTGTGAGGTGA 1067

QY 901 CCCCCTGCATGTGAGTCACTGCTGCTGCTGCGGAGGATCCGGGCGAGTGGGCGAC 960
DB 1068 CCCCCTGCATGTGAGTCACTGCTGCTGCTGCGGAGGATCCGGGCGAGTGGGCGAC 1127

QY 961 CAAGGAGCTTCTTGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTCTGTGCAAT 1020

DB 1128 CAAGGAGCTTCTTGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTCTCTGTGCAAT 1187

QY 1021 CATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACCCAGGACCGGCT 1080

DB 1188 CATGATCAGCACCTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACCCAGGACCGGCT 1247

QY 1081 TCTGGGCGGCTGAAGGGCAAGGCAAGGCAAGGCAAGGAC 1114

DB 1248 TCTGGGCGGCTGAAGGGCAAGGCAAGGCAAGGCAAGGAC 1281

RESULT 2
US-09-234-613-68
; Sequence 68, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shan, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYN00AT01
; CLONE: 724157
; US-09-234-613-68

Query Match 94.6%; Score 1054.4; DB 3; Length 1643;
Best Local Similarity 97.8%; Pred. No. 3.9e-264;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGATCAGCCCTCCAGCAAAATGGT 60
DB 192 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGGATCAGCCCTCCAGCAAAATGGT 251

QY 61 GGCCTCAGGACCGGGCTGTGTACTCTCTTCATGACACCCCTGGAGCTGTGA 120

Db 252 GGCCTCAGGACCGGGCTGGTGGTACCTCTCTCTTCATGACACCCCTGGAGCTGGTAA 311
 Qy 121 GGTTCGCTGCAGTCTCAGCGGCCCTCATGCGCAGCGAGCTGATGCTCTCTCCAGACT 180
 Db 312 GGTTCGCTGCAGTCTCAGCGGCCCTCATGCGCAGCGAGCTGATGCTCTCTCCAGACT 371
 Qy 181 GTGAGGCTCTCTATACAAATTTGGCCTCTCTCTCCAAATCCAGGAAAGTCTCTCT 240
 Db 372 GTGAGGCTCTCTATACAAATTTGGCCTCTCTCTCCAAATCCAGGAAAGTCTCTCT 407
 Qy 241 GTATTCAATGCTGCTCTGAGGCTCTGTATCTGTCGCCAATGTCGCCGCTGTCGCAC 300
 Db 408 GTATTCAATGCTGCTCTGAGGCTCTGTATCTGTCGCCAATGTCGCCGCTGTCGCAC 467
 Qy 301 CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGATGCTCTCTGGAAGATCTGTAG 360
 Db 468 CTGTTTCAAGACCTTACCCGCTTCACTGGCACCATGATGCTCTCTGGAAGATCTGTAG 527
 Qy 361 GCAGAGGACACAGGACCTCTGAGGCGCTCTCCGCGACCTCTGATGATGATCTGTGC 420
 Db 528 GCAGAGGACACAGGACCTCTGAGGCGCTCTCCGCGACCTCTGATGATGATCTGTGC 587
 Qy 421 AGCTACCGCATCTACTTCACTGCTATGACCACTGAGGCTCTCTCTGCTGCTGAGC 480
 Db 588 AGCTACCGCATCTACTTCACTGCTATGACCACTGAGGCTCTCTCTGCTGCTGAGC 647
 Qy 481 CTTGACTCTGACTCTACGACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 648 CTTGACTCTGACTCTACGACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707
 Qy 541 GACTGTGATCAGCCCTGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 708 GACTGTGATCAGCCCTGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
 Qy 601 CCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 768 CCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827
 Qy 661 GCTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
 Db 721 CTATGAGCTGGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 888 CTATGAGCTGGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 947
 Qy 781 CATGAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 948 CATGAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1007
 Qy 841 TGACGTGTAAGACCAACCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 1008 TGACGTGTAAGACCAACCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1067
 Qy 901 CCCCCTGATGTCATCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 1068 CCCCCTGATGTCATCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
 Qy 961 CAAGGACTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 1128 CAAGGACTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187
 Qy 1021 CATGATCAGACCTATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 1188 CATGATCAGACCTATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1247
 Qy 1081 TCTGGCGGCTGAAGGCGCAAGGAGGCAAGGAC 1114
 Db 1248 TCTGGCGGCTGAAGGCGCAAGGAGGCAAGGAC 1281

RESULT 3

US-09-605-785-387
 ; Sequence 387, Application US/09605785
 ; Patent No. 6321716
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqi
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darriek
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Hepler, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C16
 ; CURRENT APPLICATION NUMBER: US/09/605,785
 ; CURRENT FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 835
 ; SOFTWARE: FastSBQ for Windows Version 3.0
 ; SEQ ID NO: 387
 ; LENGTH: 537
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-605-785-387

Query Match 15.3%; Score 170; DB 4; Length 537;
 Best Local Similarity 100.0%; Pred. No. 4.3e-35;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 945 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCCTCGGATCATCAAGGCTG 1004
 Db 1 GGGCCGAGTCGGGACCAAGGAGCTCTTTCAGGCTTCTTCCCTCGGATCATCAAGGCTG 60

Qy 1005 CCCCCTCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAGCTTCTTCCAGAGC 1064
 Db 61 CCCCCTCTCTGTGCCATCATGATCAGCACCTATGATTCGGCAAAGCTTCTTCCAGAGC 120

Qy 1065 TGAACCAAGACCGGCTTCTGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
 Db 121 TGAACCAAGACCGGCTTCTGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 170

RESULT 4
 US-09-439-313-387
 ; Sequence 387, Application US/09439313
 ; Patent No. 6329505
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan Louise
 ; APPLICANT: Jiang Yuqi
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Kalos, Michael
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Retter, Mark
 ; APPLICANT: Soik, John
 ; APPLICANT: Day, Craig
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C9
 ; CURRENT APPLICATION NUMBER: US/09/439,313
 ; CURRENT FILING DATE: 1999-11-12
 ; NUMBER OF SEQ ID NOS: 575

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-439-313-387

Query Match          15.3%; Score 170; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.3e-35;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGACCAAGGACTCTTTGAGGCTTCTTCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTGAGGCTTCTTCTCGGATCATCAAGGCTG 60

QY 1005 CCCCTCTCTGTCATCATGATCATGACACCTATGATTCGGGAAAAAGCTTTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCATGACACCTATGATTCGGGAAAAAGCTTTTCCAGAGGC 120

QY 1065 TGAACCAAGGACCGCTTCTGGGCGGCTCAAGGGGCAAGGAGCAAGGAC 1114
Db 121 TGAACCAAGGACCGCTTCTGGGCGGCTCAAGGGGCAAGGAGCAAGGAC 170

RESULT 5
US-09-352-616A-387
; Sequence 387, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqui
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352.616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-352-616A-387

Query Match          15.3%; Score 170; DB 4; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.3e-35;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGACCAAGGACTCTTTGAGGCTTCTTCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTGAGGCTTCTTCTCGGATCATCAAGGCTG 60

QY 1005 CCCCTCTCTGTCATCATGATCATGACACCTATGATTCGGGAAAAAGCTTTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCATGACACCTATGATTCGGGAAAAAGCTTTTCCAGAGGC 120

QY 1065 TGAACCAAGGACCGCTTCTGGGCGGCTCAAGGGGCAAGGAGCAAGGAC 1114
Db 121 TGAACCAAGGACCGCTTCTGGGCGGCTCAAGGGGCAAGGAGCAAGGAC 170

RESULT 6
US-09-188-930-23
; Sequence 23, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
```

```
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 997
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-23

Query Match          5.5%; Score 61; DB 3; Length 997;
Best Local Similarity 52.1%; Pred. No. 9.4e-07;
Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 313 CCTACCCGCTTCACTGGCACCATGGATGCCCTTCGTGAAGATCGTGAGGCACGAGGCAC 372
Db 718 CCGCAGCAACAACATGTCATGTAGGTGGATTACACAGATGATTCGAGAAGGGGAGC 777

QY 373 CAGAACCTCTGGAGCGGCTCCCGCCACCCCTGGTGTGATGACTGTGCCAGCTACCGCCAT 432
Db 778 CAAGTCACCTCTGGCGGGCAACGGCATCAATGTCTCAAAATTTGCCCTGAGTCGGCCAT 837

QY 433 CTACTTCACCTGCCTATGACCAACTCAAGGCGCTTCCTGTGTGTGGAGCCCTGACTCTGA 492
Db 838 CAATTCATGGCAATGACAGATGAACGGCTTTGCGGTAGTGTACAGGAGAGCGCTGAG 897

QY 493 CCTCTACGCCACCCATCGTGGCTGGCGCGCTGGCCGCGCTGGGACCGTGTGATGATCAG 552
Db 898 GATCCACGAAAGGCTTTGTGGCAGGCTCCTTTGGCGGAGCCATTGGCCAGAGTAGCATCTA 957

QY 553 CCCCTGGAGCTTATGGGAC 573
Db 958 CCCAATGGAGGTTCTGAAGAC 978

RESULT 7
US-09-188-930-262
; Sequence 262, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 262
; LENGTH: 1816
; TYPE: DNA
; ORGANISM: Mouse
US-09-188-930-262

Query Match          5.5%; Score 61; DB 3; Length 1816;
Best Local Similarity 52.1%; Pred. No. 1.2e-06;
Matches 136; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 313 CCTACCCGCTTCACTGGCACCATGGATGCCCTTCGTGAAGATCGTGAGGCACGAGGCAC 372
Db 718 CCGCAGCAACAACATGTCATGTAGGTGGATTACACAGATGATTCGAGAAGGGGAGC 777

QY 373 CAGGACCTCTGGAGCGGCTCCCGCCACCCCTGGTGTGATGACTGTGCCAGCTACCGCCAT 432
Db 778 CAAGTCACCTCTGGCGGGCAACGGCATCAATGTCTCAAAATTTGCCCTGAGTCGGCCAT 837
```

QY 433 CTACTTCAGTCTGCTATGACCACTGAAGGCTTCTCTGTGTGAGGCTTGAACCTCTGA 492
 Db 838 CAAATTCATGGCATATGACGATGAAACCGCTTGTGTGAGTATGATCAGGAGACGCTGAG 897
 QY 493 CCTCTAGCAGCACCATGGTGGCTGGCGGCTTGGCCGCTTGGGACCGGTGACTGTGATCAG 552
 Db 898 GATCCAGGAAGGCTTGTGGAGGCTCTTGGCGGAGCCATTGCCAGAGTAGCATCTA 957
 QY 553 CCCCTGGAGCTTATGCGGAC 573
 Db 958 CCAATGGAGTTCTGAAGAC 978

RESULT 8

US-09-068-140A-9
 ; Sequence 9, Application US/09068140A
 ; Patent No. 6281409
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
 ; PPLICANT: and Rex Michael Brennan
 ; TITLE OF INVENTION: Blackcurrant Promoters and Genes
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-0939
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/068,140A
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/EP96/04807
 ; FILING DATE: No. 6281409ember 4, 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dinner, Dara L.
 ; REGISTRATION NUMBER: 33,680
 ; REFERENCE/DOCKET NUMBER: C70237
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-270-5017
 ; TELEFAX: 610-270-5090
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1311 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Ribes nigrum
 ; STRAIN: Ben Alder
 ; US-09-068-140A-9

Query Match 4.4%; Score 48.8; DB 4; Length 1311;
 Best Local Similarity 45.7%; Pred. No. 0.0015;
 Matches 253; Conservative 0; Mismatches 292; Indels 9; Gaps 2;
 QY 318 CCCGTTTCACCTGGCACCATGATGCTTCCTGTAAGATCGTGAGGCGACGAGCCACGGA 377
 Db 206 CTCATATCGCGGCTCTCCGACAGCCCTTGGGTGATGATACTGAAAGTTGAAGTCCCGCG 265
 QY 378 CCTCTGGAGCGGCTCCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCCATCTACT 437

Db 266 GACITTTACCGTGGCAATGGTGCATTCGGTCTCGGTGCGAGGACGCTCAGCAGTGATT 325
 QY 438 TCACCTGCTATGACCACTGAAGGCTTCTCTGTGTGAGGCTTGAACCTCTGACCTCT 497
 Db 326 TCTCCGTTACGAGATGTGTAAGGAGACTTTTCTCATGTGTATCCGAGC---AATTCCG 382
 QY 498 AGCACCCATGTGTGGCTGGCGGCTGGCCGCTGGGACCGGTGATCTGTGATCAGCCCC 557
 Db 383 GTGCGCAGCCGTTTCGGGGGTGTTCCGGAGCGGTGGCAAGCGCGGTGATTACGCCGA 442
 QY 558 TGGAGCTTATGCGGCAAAAGCTGCGAGGCTCAGCATGTGTCGTACCGGAGCTGGGTGCT 617
 Db 443 TGGATGTGTGTAAGAACAGAGTTGTCAGTTGTCAGAGCAGTCCGTACAGGGGTGTTGTTGATT 502
 QY 618 GTGTTTGAACCTGAGTGGCTCAGGCTGGGCTGGGCTGCTCAGTGTGGTGGGCTGGGCCCCA 677
 Db 503 GCGTGGAGGGGTGTTGGTAGAAGAGGATTGGCGCATTTTACGCATCTTATCGAACAA 562
 QY 678 CTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCACTATGAGCTGGTGAAGA 737
 Db 563 CTGTGTCATGATGCCCGTTTAGCGCCGTTCACTTCCCCACATATGAAGCCACGAAGA 622
 QY 738 GCTGCTCAATGGGCTCAGGCCGAA-----GGACCAAGACTTCTGTGGGCAATGAGCTTTG 791
 Db 623 AAGGCTTGTGGAGGTGTGCGCGGAGACTCGGAACGATGAGAATTTGTTAGTGCATGCTA 682
 QY 792 TGGCTGGTGCATCTCAGGAGCGGTGGCTCAGTGTGCTCAGTGTGCTTACCCCTTTGACGTGTAA 851
 Db 683 CTGCTGGTCTGCTCTGTGAGCTTTGGCTGCTAGTAGTAACCACTCCACTAGATGTTGCTCA 742
 QY 852 AGACCCCAACGCCAG 865
 Db 743 AACTCAGTTGCAG 756

RESULT 9

US-09-068-140A-14
 ; Sequence 14, Application US/09068140A
 ; Patent No. 6281409
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
 ; APPLICANT: and Rex Michael Brennan
 ; TITLE OF INVENTION: Blackcurrant Promoters and Genes
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-0939
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/068,140A
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/EP96/04807
 ; FILING DATE: No. 6281409ember 4, 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dinner, Dara L.
 ; REGISTRATION NUMBER: 33,680
 ; REFERENCE/DOCKET NUMBER: C70237
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-270-5017
 ; TELEFAX: 610-270-5090
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1311 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Ribes nigrum
 ; STRAIN: Ben Alder
 ; US-09-068-140A-9


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 5150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; 'S-09-068-140A-14

Query Match          4.4%; Score 48.8; DB 4; Length 5150;
Best Local Similarity 45.7%; Pred. No. 0.0024;
Matches 253; Conservative 0; Mismatches 292; Indels 9; Gaps 2;

QY 318 CCGCTTCACTGGCACCATGGATCGCTTCTGAAGATCGTGAGGACGAGGACGAGCA 377
DB 3370 CTCATCCGCGGTCTCGACAGCCCTTGGTCGATCTGAAGTTGAAGTCCCGCG 3429

QY 378 CCCTCTGGACGGCCCTCCCGCCGCTGCTGATGATGCTGCGACCTACCCCATCTACT 437
DB 3430 GACTTTACCGTGGCATTTGGTGAATGGTCTCGGTGACGACCACTCACGACGTATT 3489

QY 438 TCATGCTCTATGACCACTGAAGGCTTCTGTTGTTGTTGAGCCCTGACCTTCACTCT 497
DB 3490 TCTCCGTTTACGAGATGTGTAAAGGAGCTTTTCTCATGCTGATCCGAGC---AATTCCG 3546

QY 498 ACGCACCATGTTGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGG 557
DB 3547 GTGGCGACGGCTTTCCGGGGTGTTCGGGACGCTGGCGGCTGGCGGCTGGCGGCTGGCGG 3606

QY 558 TGGAGCTTATCGGCAAAAGCTGCAAGCTGATGCTGCTACCGGAGCTGGGTCCT 617
DB 3607 TGGATGTTGTAACAGAGGTTTGCAGTTGTCAGAGCAGCTCGGTACAAAGGCTGTTGTTGATT 3666

QY 618 GTGTTGAACTGCACTGGCTCAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
DB 3667 CGCTGAGGAGGCTGTTGTTGTAAGAGGAGTTGGCGCATTTACGATCTTATCGAACAA 3726

QY 678 CTGCGCTTCAGATGTCCTCTTCTCAGCCCTGTTACTGGTTCAACTATGAGTGGTGAAGA 737
DB 3727 CTGTGTCATGATGCGCGGTTTACGCGGCTTCACTTCGCCCAATATGAGCCACGAGA 3786

QY 738 GCTGGCTCAATGGGCTCAGCCGAA-----GACACAGATTCTGTGGGCATGAGCTTTG 791
DB 3787 AAGGCTTGTGGAGGTGTCGCGGAGACTGCGGAAAGATTTGTTAGTGCATGCTA 3846

QY 792 TGGCTGTTGCTATCTCAGGACGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
DB 3847 CTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3906

QY 852 AGACCAACGCCAG 865
DB 3907 AAACCTAGTTCAG 3920

```

RESULT 10

```

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA

```

```

; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

```

Query Match 4.3%; Score 47.6; DB 1; Length 7218;

Best Local Similarity 5.0%; Pred. No. 0.0055;

Matches 20; Conservative 213; Mismatches 167; Indels 0; Gaps 0;

```

QY 132 AGTCTACGGCCCTCCATGGCCAGGAGCTGATGCTTCTCCACACTGTGGAGCTCT 191
DB 1030 ATTCCGAGCTTGGCTCGAGTGGAGGAGCTTGGATYYYYYYYYYYYYYYYYYY 1089

QY 192 CCTATACAAATTGCCCTCTCTCTCCCAATCCACAGGAAAGTCCCTCTGTATTCGAATG 251
DB 1090 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1149

QY 252 GTGCTCTGGAGCCTGTGTACCTGTGCCCAATGGTCCCGCTGTGCCACCTGGTTCAAG 311
DB 1150 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1209

QY 312 ACCCTACCGCTTCACTGSCACCATGGATGCTTCTGTAAGATCGTGAGGACGAGGCA 371
DB 1210 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1269

QY 372 CCAGGACCTCTGGAGCGGCTCCCGCCACCTGGTGTGATGACTGTCCAGCTACCGCA 431
DB 1270 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1329

QY 432 TCTACTTCACTGCTATGACCAACTGAAGGCTTCTCTGTGTGTCGAGCCCTGACCTGTG 491
DB 1330 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1389

QY 492 ACCTCTACGACCACTGGTGGCTGGCGCTGGCCCGCT 531
DB 1390 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1429

```

RESULT 11

```

US-08-933-750C-61
; Sequence 61, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti

```

```

; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNOT02
; CLONE: 207452
; US-08-933-750C-61

Query Match 3.8%; Score 42.2; DB 2; Length 1594;
Best Local Similarity 59.7%; Pred. NO. 0.082;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 951 AGTCGGGCACCAAGGACTCTTTGCGAGGCTTCCTTCGCGATCATCAAGGCTGCCCT 1010
Db 1001 AGAAGGCGCCCTGGGCTCTTCAAGGCGCTGCCAGCTTGCTGAAGGCTGCCCTCT 1060

QY 1011 CTGTGCGCATCATGATCAGCACTATGAGTTGGCAAAAGCTTCTCCAGAGGCTGAAC 1069
Db 1061 CCACAGGCTTCATGTTCTTCTCGATGAATCTTCTGTAATGTTCTTCCACTGCATGAAC 1119

RESULT 12
US-09-234-613-61
; Sequence 61, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.

```

```

; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNOT02
; CLONE: 207452
; US-09-234-613-61

Query Match 3.8%; Score 42.2; DB 3; Length 1594;
Best Local Similarity 59.7%; Pred. NO. 0.082;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 951 AGTCGGGCACCAAGGACTCTTTGCGAGGCTTCCTTCGCGATCATCAAGGCTGCCCT 1010
Db 1001 AGAAGGCGCCCTGGGCTCTTCAAGGCGCTGCCAGCTTGCTGAAGGCTGCCCTCT 1060

QY 1011 CTGTGCGCATCATGATCAGCACTATGAGTTGGCAAAAGCTTCTCCAGAGGCTGAAC 1069
Db 1061 CCACAGGCTTCATGTTCTTCTCGATGAATCTTCTGTAATGTTCTTCCACTGCATGAAC 1119

RESULT 13
US-08-961-871-11
; Sequence 11, Application US/08961871
; Patent No. 6013858
; GENERAL INFORMATION:
; APPLICANT: Wallace, Douglas C.
; APPLICANT: Graham, Brett H.
; APPLICANT: MacGregor, Grant R.
; TITLE OF INVENTION: Mouse Lacking Heart-Muscle Adenine
; Nucleotide Translocator Protein and Methods
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

Db 415 GCGGTCTCGAGGACACCCGCTCAGCGCGTCTCTCCACCTGCGCGCCCAACCGTGGACTCC 474
Qy 491 GACCTCTACGCAACCCATGGTGGCGCGCTGGCCCGCCTGGGCAACCGTGGACTGTGATC 550
Db 475 GAGCGGCTCGCGCGCACCGACCCGGAGGCACTGCGCCCGGCTCGTAACCGGGAAGGCCACC 534
Qy 551 AGCCCCCTGGAGCT 564
Db 535 GCCGCGCTGCACCT 548

```

Search completed: January 22, 2003, 19:05:41
 Job time : 75.6489 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 17:33:27 ; Search time 55.7 Seconds
(without alignments)
8985.296 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 ctgaagcttcaagatggctg.....aggggcaaggaggcaagac 1114

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 395772 seqs, 224632407 residues
Tr number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1114	100.0	1114	10	US-09-888-358-1
2	1054.4	94.6	1546	10	US-09-888-358-2
3	1054.4	94.6	1643	10	US-09-840-787-68
4	185.8	16.7	453	10	US-09-867-701-4486
5	170.2	15.3	512	10	US-09-917-800A-329
6	170.2	15.3	537	9	US-10-012-886-387
7	170.2	15.3	537	9	US-09-895-793-387
8	170.2	15.3	537	9	US-09-895-814-387
9	170.2	15.3	537	10	US-09-759-143-387
10	170.2	15.3	537	10	US-09-780-669-387
11	170.2	15.3	537	10	US-09-822-827-387
12	154	13.8	275	10	US-09-908-711-46
13	116.2	10.4	246	10	US-09-960-352-2543
14	93.4	8.4	452	10	US-09-880-107-1133
15	74.8	6.7	452	10	US-09-864-761-11674
16	74.8	6.7	496	10	US-09-864-761-11708
17	60	5.4	3334	9	US-09-992-598-288
18	60	5.4	3334	9	US-09-989-293A-288
19	60	5.4	3334	9	US-10-063-547-57

20	60	5.4	3334	9	US-09-989-735-288	Sequence 288, App
21	60	5.4	3334	9	US-09-990-444-288	Sequence 288, App
22	60	5.4	3334	9	US-09-989-730-288	Sequence 288, App
23	60	5.4	3334	9	US-09-990-436-288	Sequence 288, App
24	60	5.4	3334	9	US-09-991-181-288	Sequence 288, App
25	60	5.4	3334	9	US-09-993-687-288	Sequence 288, App
26	60	5.4	3334	9	US-09-989-734-288	Sequence 288, App
27	60	5.4	3334	9	US-09-997-653-288	Sequence 288, App
28	60	5.4	3334	9	US-10-063-616-57	Sequence 57, Appl
29	60	5.4	3334	10	US-09-989-722-288	Sequence 288, App
30	60	5.4	3334	10	US-09-989-723-288	Sequence 288, App
31	60	5.4	3334	10	US-09-989-279-288	Sequence 288, App
32	60	5.4	3334	10	US-09-989-727-288	Sequence 288, App
33	60	5.4	3334	10	US-09-989-731-288	Sequence 288, App
34	60	5.4	3334	10	US-09-989-732-288	Sequence 288, App
35	60	5.4	3334	10	US-09-991-073-288	Sequence 288, App
36	60	5.4	3334	10	US-09-990-442-288	Sequence 288, App
37	60	5.4	3334	10	US-09-991-163-288	Sequence 288, App
38	60	5.4	3334	10	US-09-993-604-288	Sequence 288, App
39	60	5.4	3334	10	US-09-990-456-288	Sequence 288, App
40	60	5.4	3334	10	US-09-989-721-288	Sequence 288, App
41	60	5.4	3334	12	US-10-006-867-57	Sequence 57, Appl
c 42	57.6	5.2	144	10	US-09-864-761-28253	Sequence 28253, A
44	52.2	4.7	3555	12	US-10-044-090-527	Sequence 527, App
45	48.8	4.4	1311	10	US-09-800-528-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-888-358-1
; Sequence 1, Application US/09888358
; Patent No. US20020119137A1
; GENERAL INFORMATION:
; APPLICANT: LEWIN, DAVID
; APPLICANT: ADAMS, SEAN H.
; APPLICANT: YU, XING XIAN
; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 10716/66
; CURRENT APPLICATION NUMBER: US/09/888,358
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,307
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-358-1

Query Match	100.0%	Score 1114;	DB 10;	Length 1114;
Best Local Similarity	100.0%	Pred. No. 7e-286;		
Matches 1114;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	CTCAAGCTTCAAGATGGCTGACAGGACCCCTCGGGGATCAGCCCCCTCCAGCAAAATGGT	60		
Db 1	CTCAAGCTTCAAGATGGCTGACAGGACCCCTCGGGGATCAGCCCCCTCCAGCAAAATGGT	60		
Qy 61	GGCTCAGGACCCGGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGACCTGGTGA	120		
Db 61	GGCTCAGGACCCGGGGCTGTGGTTACCTCTCTCTTCATGACACCCCTGACCTGGTGA	120		
Qy 121	GTTTCGCTCAGTCTCAGCGGCCCTCCATGCGCAGCTGATGCTTCCTCCAGACT	180		
Db 121	GTTTCGCTCAGTCTCAGCGGCCCTCCATGCGCAGCTGATGCTTCCTCCAGACT	180		
Qy 181	GTGAGCTCTCTCTATACCAAAATGGCTCTCTCTCTCAATCCACGGAAGTGCCTCT	240		
Db 181	GTGAGCTCTCTCTATACCAAAATGGCTCTCTCTCTCAATCCACGGAAGTGCCTCT	240		

QY 241 GTATTGAATGGTCTCTGAGGCTCTGACTGTGCGCAATGGTGGCCGCTGTGCAC 300
 Db 241 GTATTGAATGGTCTCTGAGGCTCTGACTGTGCGCAATGGTGGCCGCTGTGCAC 300
 QY 301 CTGGTTTCAAGACCCCTACCGCTTCTACCTGGCAACATGATGCTTCTGTAAGATCGTGAG 360
 Db 301 CTGGTTTCAAGACCCCTACCGCTTCTACCTGGCAACATGATGCTTCTGTAAGATCGTGAG 360
 QY 361 GCACGAGGACACAGGACCCCTCTGGAGGGCTTCCCGCCACCCCTGGTGTGACTGTGCC 420
 Db 361 GCACGAGGACACAGGACCCCTCTGGAGGGCTTCCCGCCACCCCTGGTGTGACTGTGCC 420
 QY 421 AGCTACCGCATCTACTTCTACCTGCTATGACCACTGAAGGCTTCTGCTGTGCTGAGC 480
 Db 421 AGCTACCGCATCTACTTCTACCTGCTATGACCACTGAAGGCTTCTGCTGTGCTGAGC 480
 QY 481 CTTGACCTCTGACCTCTACCGCACCCATGGTGGCTGGCGGCTGGCCGCTGGGACCGT 540
 Db 481 CTTGACCTCTGACCTCTACCGCACCCATGGTGGCTGGCGGCTGGCCGCTGGGACCGT 540
 QY 541 GACTGTGATCAGCCCTCTGAGGCTTATGCGGACAAAGCTGCAAGCTCAGGCTGAGTGTGCTA 600
 Db 541 GACTGTGATCAGCCCTCTGAGGCTTATGCGGACAAAGCTGCAAGCTCAGGCTGAGTGTGCTA 600
 QY 601 CCGGAGCTGGTGGCTGTGTTGCACTGCTAGTGGCTCAGGCTGGCTGGCTCAGTGTG 660
 Db 601 CCGGAGCTGGTGGCTGTGTTGCACTGCTAGTGGCTCAGGCTGGCTGGCTCAGTGTG 660
 QY 661 GCTGGGCTGGGCCCCCAGCTGCTGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
 Db 661 GCTGGGCTGGGCCCCCAGCTGCTGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
 QY 721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCTGAGGCTGAGTGTGTTGGG 780
 Db 721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCTGAGGCTGAGTGTGTTGGG 780
 QY 781 CATGACTTTTGGCTGGTGGCTCTCAGGACCGTGGCTGAGTGGCTGACTTACCTT 840
 Db 781 CATGACTTTTGGCTGGTGGCTCTCAGGACCGTGGCTGAGTGGCTGACTTACCTT 840
 QY 841 TGAGTGGTAAAGACCCAGCCAGTGGCTCTGCGAGCGATGGAGGCTGTGAGAGTGA 900
 Db 841 TGAGTGGTAAAGACCCAGCCAGTGGCTCTGCGAGCGATGGAGGCTGTGAGAGTGA 900
 QY 901 CCCCTGCATGTGGAATCCACTGCTGCTGCGGAGGATCCGCGCGAGTGGGAC 960
 Db 901 CCCCTGCATGTGGAATCCACTGCTGCTGCGGAGGATCCGCGCGAGTGGGAC 960
 QY 961 CAAGGACTCTTGGCAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTTGCCAT 1020
 Db 961 CAAGGACTCTTGGCAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTGTTGCCAT 1020
 QY 1021 CATGATCAGCACCTATGAGTTCGGCAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1080
 Db 1021 CATGATCAGCACCTATGAGTTCGGCAAGCTTCTTCCAGAGGCTGAACAGGACCGGCT 1080
 QY 1081 TCTGGGCGGTGAAGAGGCGAAGAGGCAAGGAC 1114
 Db 1081 TCTGGGCGGTGAAGAGGCGAAGAGGCAAGGAC 1114

RESULT 2

US-09-888-358-2
 ; Sequence 2, Application US/09888358
 ; Patent No. US20020119137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LEWIN, DAVID
 ; APPLICANT: ADAMS, SEAN H.
 ; APPLICANT: YU, XING XIAN
 ; TITLE OF INVENTION: CGI-69 COMPOSITIONS AND METHODS OF USE
 ; FILE REFERENCE: 10716/56
 ; CURRENT APPLICATION NUMBER: US/09/888,358
 ; CURRENT FILING DATE: 2001-06-22

; PRIOR APPLICATION NUMBER: 60/213,307
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1546
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-888-358-2
 Query Match 94.6%; Score 1054.4; DB 10; Length 1546;
 Best Local Similarity 97.8%; Pred. No. 4.9e-270;
 Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
 QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCTTGGGGCATCAGCCCCCTCCAGCAAAATGGT 60
 Db 105 CTGAAGCTTCAAGATGGCTGACAGGACCTTGGGGCATCAGCCCCCTCCAGCAAAATGGT 164
 QY 61 GGCCTCAGGACACCGGGCTGTGGTTACCTCTCTTCATGACACCCCTGGAGCTGTGAA 120
 Db 165 GGCCTCAGGACACCGGGCTGTGGTTACCTCTCTTCATGACACCCCTGGAGCTGTGAA 224
 QY 121 GGTTCGGCTGCACTCTCAGCGCCCTCCATGGCCAGCGAGCTGATGCCCTTCTCCAGACT 180
 Db 225 GGTTCGGCTGCACTCTCAGCGCCCTCCATGGCCAGCGAGCTGATGCCCTTCTCCAGACT 284
 QY 181 GTGAGGCTCTCTCTTACCAAAATTCGCCCTCTCTCTCAATCCACAGGGAAGTGCCTCT 240
 Db 285 GTGAGGCTCTCTCTTACCAAAATTCGCCCTCTCTCTCAATCCACAGGGAAGTGCCTCT 320
 QY 241 GTATGCAATGGTGTCTCTGGAGCTCTGTACTGTGCCCCAAATGGTGGCCGCTGTGCAC 300
 Db 321 GTATGCAATGGTGTCTCTGGAGCTCTGTACTGTGCCCCAAATGGTGGCCGCTGTGCAC 380
 QY 301 CTGGTTTCAAGACCCCTACCGCTTCTACCTGGCACCATGGATGCTTCTGTAAGATCGTGAG 360
 Db 381 CTGGTTTCAAGACCCCTACCGCTTCTACCTGGCACCATGGATGCTTCTGTAAGATCGTGAG 440
 QY 361 GCACGAGGACACAGGACCCCTCTGAGGCGGCTTCCCGCCACCCCTGGTGTGACTGTGCC 420
 Db 441 GCACGAGGACACAGGACCCCTCTGAGGCGGCTTCCCGCCACCCCTGGTGTGACTGTGCC 500
 QY 421 AGCTACCGCATCTACTTCTACCTGCTATGACCACTGAAGGCTTCTGTTGGTGGAGC 480
 Db 501 AGCTACCGCATCTACTTCTACCTGCTATGACCACTGAAGGCTTCTGTTGGTGGAGC 560
 QY 481 CTTGACCTCTGACTCTTACGCGACCCATGGTGGCTGGCGGCTGGCCGCTGGGACCGT 540
 Db 561 CTTGACCTCTGACTCTTACGCGACCCATGGTGGCTGGCGGCTGGCCGCTGGGACCGT 620
 QY 541 GACTGTGATCAGCCCCCTGGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTGCTA 600
 Db 621 GACTGTGATCAGCCCCCTGGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTGCTA 680
 QY 601 CCGGAGCTGGTGGCTGTGTTGCACTGCTGAGTGGCTCAGGCTGGCTGGCTCAGTGTG 660
 Db 681 CCGGAGCTGGTGGCTGTGTTGCACTGCTGAGTGGCTCAGGCTGGCTGGCTCAGTGTG 740
 QY 661 GGTGGCTGGGCGCCCACTGCTTCCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
 Db 741 GGTGGCTGGGCGCCCACTGCTTCCAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 800
 QY 721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCCGAGGACGAGCTTCTGTGGG 780
 Db 801 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCCGAGGACGAGCTTCTGTGGG 860
 QY 781 CATGAGCTTTTGGCTGGTGGCTCTCAGGACCGTGGCTGCTGAGTGTGACTTACCTT 840
 Db 861 CATGAGCTTTTGGCTGGTGGCTCTCAGGACCGTGGCTGCTGAGTGTGACTTACCTT 920
 QY 841 TGACCTGTGAAGACCCAGGCGAGGCTGGCTCTGGGAGCGATGGAGGCTGTGAGAGTGA 900
 Db 921 TGACCTGTGAAGACCCAGGCGAGGCTGGCTCTGGGAGCGATGGAGGCTGTGAGAGTGA 980

QY 901 CCCCTGATGTGACTCCACCTGCTGCTGCGAGGATCCGGCGAGTGGGCAC 960
Db 981 CCCCTGATGTGACTCCACCTGCTGCTGCGAGGATCCGGCGAGTGGGCAC 1040
QY 961 CAAGGACTCTTTGAGGCTTCTTCCGAGATCATCAAGGCTGCCCCCTCTCTGTCAT 1020
Db 1041 CAAGGACTCTTTGAGGCTTCTTCTCGGATCATCAAGGCTGCCCCCTCTCTGTCAT 1100
QY 1021 CATGATCAGACACTATGAGTTCGGCAAGCTTCTTCCAGGCTCAACGAGGACCGGCT 1080
Db 1101 CATGATCAGACACTATGAGTTCGGCAAGCTTCTTCCAGGCTCAACGAGGACCGGCT 1160
QY 1081 TCTGGGCGCTGAAAGGCGCAAGGAGGCAAGGAC 1114
Db 1161 TCTGGGCGCTGAAAGGCGCAAGGAGGCAAGGAC 1194

RESULT 3

US-840-787-68
ence 68, Application US/09840787
ent No. US20020059264A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
Hillman, Jennifer L.
Bandman, Olga
Shah, Purvi
Au-Young, Janice
Yue, Henry
Guegler, Karl J.
Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/840,787
FILING DATE: 23-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/518,865
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYN0A01
CLONE: 724157
SEQUENCE DESCRIPTION: SEQ ID NO: 68 :
US-09-840-787-68

Query Match

94.6%; Score 1054.4; DB 10; Length 1643;

Best Local Similarity 97.8%; Pred No. 5e-270;
Matches 1089; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 1 CTGAAGCTTCAAGATGGCTGACACGAGACCTTGCAGGATCAGCCCCCTCCAGCAAAATGGT 60
Db 192 CTGAAGCTTCAAGATGGCTGACACGAGACCTTGCAGGATCAGCCCCCTCCAGCAAAATGGT 251
QY 61 GGCCTCAGGACACGGGGCTGTGGTTACCTCTCTTTCATGACACCTCTGAGCTGGTAA 120
Db 252 GGCCTCAGGACACGGGGCTGTGGTTACCTCTCTTTCATGACACCTCTGAGCTGGTAA 311
QY 121 GGTTCGCTCGAGTCTCAGGGCCCTCCATGGCCAGCAGCTGATCCCTTCTCCAGACT 180
Db 312 GGTTCGCTCGAGTCTCAGGGCCCTCCATGGCCAGCAGCTGATCCCTTCTCCAGACT 371
QY 181 GTGAGGCTCTTCTTATACCAAAATTCCTCTCTCCAAATCCACAGGGAAGTGGCTCCT 240
Db 372 GTGAGGCTCTTCTTATACCAAAATTCCTCTCTCCAAATTCCTCTCTCCAAATTCCTCTCT 407
QY 241 GTATGCAATGGTGTCTCTGGAGGCTCTGTACTCTGCGCCAAATGGTGGCCGTGTGCCAC 300
Db 408 GTATGCAATGGTGTCTCTGGAGGCTCTGTACTCTGCGCCAAATGGTGGCCGTGTGCCAC 467
QY 301 CTGGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCTTCTGTAAGATCGTGAG 360
Db 468 CTGGTTTCAAGACCTTACCCGCTTCACTGGCACCATGGATGCTTCTGTAAGATCGTGAG 527
QY 361 GCACGAGGACACGAGACCTCTGGAGCGGCTTCCCGCACCTCTGGTGTGATGCTGTGCC 420
Db 528 GCACGAGGACACGAGACCTCTGGAGCGGCTTCCCGCACCTCTGGTGTGATGCTGTGCC 587
QY 421 AGCTACCGCATCTACTTCACTGCTATGACCAATGGAAGGCTTCTCTGTGTGTCAGC 480
Db 588 AGCTACCGCATCTACTTCACTGCTATGACCAATGGAAGGCTTCTCTGTGTGTCAGC 647
QY 481 CCTGACCTCTGACCTCTACGACACCATGGTGGCTGGCGCTGGCCCGCTGGGACCGT 540
Db 648 CCTGACCTCTGACCTCTACGACACCATGGTGGCTGGCGCTGGCCCGCTGGGACCGT 707
QY 541 GACTGTATCAGCCCCCTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGTGCTA 600
Db 708 GACTGTATCAGCCCCCTGGAGCTTATGCGGACAAAGCTGCAGGCTCAGCATGTGTGCTA 767
QY 601 CCGGAGCTGGGTGCTGTGTCGAACTGCACTGAGTGGCTCAGGGTGGCTGGCGCTCAGTGTG 660
Db 768 CCGGAGCTGGGTGCTGTGTCGAACTGCACTGAGTGGCTCAGGGTGGCTGGCGCTCAGTGTG 827
QY 661 GCTGGGCTGGGGCCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAA 720
Db 828 GCTGGGCTGGGGCCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTGTTCAA 887
QY 721 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCCAAGAGCACAGACTTCTGTGGG 780
Db 888 CTATGAGCTGGTGAAGAGTGGCTCAATGGGCTCAGGCCAAGAGCACAGACTTCTGTGGG 947
QY 781 CATGAGCTTTGGCTGGTGGCATCTCAGGAGCGGTGGCTGAGTGTGCTTACCTTACCTT 840
Db 948 CATGAGCTTTGGCTGGTGGCATCTCAGGAGCGGTGGCTGAGTGTGCTTACCTTACCTT 1007
QY 841 TGACGTGGTAAAGACCCCAACGCTGCTCTGGGAGCGATGGAGGCTGTGAGAGTGA 900
Db 1008 TGACGTGGTAAAGACCCCAACGCTGCTCTGGGAGCGATGGAGGCTGTGAGAGTGA 1067
QY 901 CCCCCTGCAATGTGAGACTCCACTGGCTGTGTGCGGAGGATCCGGGCGAGTGGGCGAC 960
Db 1068 CCCCCTGCAATGTGAGACTCCACTGGCTGTGTGCGGAGGATCCGGGCGAGTGGGCGAC 1127
QY 961 CAAGGACTCTTTTTCAGGCTTCTCTCTCGGATCATCAAGGCTGCCCTCTCTGTGCCAT 1020
Db 1128 CAAGGACTCTTTTTCAGGCTTCTCTCTCGGATCATCAAGGCTGCCCTCTCTGTGCCAT 1187
QY 1021 CATGATCAGACACTATGAGTTCGGCAAGCTTCTTCCAGGCTCAACGAGGACCGGCT 1080

Db 1188 CATGATCAGCCTATGAGTTCGGCAAAAGCTTTCCAGAGCTGAACAGGACCGGCT 1247

Qy 1081 TCTGGCGGCTGAAAGGGGCAAGGCAAGGAC 1114
 Db 1248 TCTGGCGGCTGAAAGGGGCAAGGCAAGGAC 1281

RESULT 4

US-09-867-701-4486
 ; Sequence 4486, Application US/09867701
 ; Patent No. US20020132237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agiate, Paul A.
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.497
 ; CURRENT APPLICATION NUMBER: US/09/867,701
 ; CURRENT FILING DATE: 2001-05-29
 ; NUMBER OF SEQ ID NOS: 10912
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4486
 ; LENGTH: 453
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-867-701-4486

Query Match 16.7%; Score 185.8; DB 10; Length 453;
 Best Local Similarity 92.0%; Pred. No. 4.3e-40;
 Matches 207; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
 Qy 503 CCCATGTTGGCTGGCGGCTGGCCGCTGGCCACCGTGTGATCAGCCCCCTGGAG 562
 Db 230 CCCATGTTGACCCCAATCCACACAGTGGGACCGTGTGATCAGCCCCCTGGAG 289
 Qy 563 CTTATGGCGCAAAAGCTGACGCTGACGATGTGTGCTACCGGAGTGGTGCCTGTGTT 622
 Db 290 CTTATGGCGCAAAAGCTGACGCTGACGATGTGTGCTACCGGAGTGGTGCCTGTGTT 349
 Qy 623 CGAAGTCAGGCTGAGGCTGGCGCTGACGCTGACGCTGAGGCTGGGCGCCACTGCC 682
 Db 350 CGAAGTCAGGCTGAGGCTGGCGCTGACGCTGAGGCTGGGCGCCACTGCC 408
 Qy 683 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 727
 Db 409 CTTGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAG 453

RESULT 5

US-09-917-800A-329/c
 ; Sequence 329, Application US/09917800A
 ; Patent No. US20020119462A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark
 ; APPLICANT: Johnson, Kory
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5038-US
 ; CURRENT APPLICATION NUMBER: US/09/917,800A
 ; CURRENT FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,040
 ; PRIOR FILING DATE: 2000-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,880
 ; PRIOR FILING DATE: 2000-11-02
 ; PRIOR APPLICATION NUMBER: US 60/290,029
 ; PRIOR FILING DATE: 2001-05-11
 ; PRIOR APPLICATION NUMBER: US 60/290,645
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: US 60/292,336

; PRIOR FILING DATE: 2001-05-22
 ; PRIOR APPLICATION NUMBER: US 60/295,798
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 60/297,457
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,884
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: US 60/303,459
 ; PRIOR FILING DATE: 2001-07-09
 ; NUMBER OF SEQ ID NOS: 1740
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 329
 ; LENGTH: 512
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA926365
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(512)
 ; OTHER INFORMATION: n = a o r c o r t
 ; US-09-917-800A-329

Query Match 15.3%; Score 170.2; DB 10; Length 512;
 Best Local Similarity 81.7%; Pred. No. 6.1e-36;
 Matches 196; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
 Qy 857 CAACGCCAGTCTGCTGGGAGCGATGAGGCTGTGAGAGTGAACCCCTGCATGTGAC 916
 Db 511 CAGCAGCAAAATGCTCAGTGGGAGCGATGAGAGTGAAGCGCCCAAGATTGAC 452
 Qy 917 TCACCTCGCTGTGCTGGGAGGATCGGGCCGAGTGGGGCCACCAAGGACTCTTTGCA 976
 Db 451 TCACCTCGCTGTGCTGGGAGATCCAGGCTGAATCTGGCACCAGGGGACTCTTTGCA 392
 Qy 977 GCTTCTCTCTCGGATCATCAAGGCTGCCCCCTCTCTGTCATCATGATCAGACCTAT 1036
 Db 391 GGTTCCTCCNCAGGATCATCAAGGCTGGCCCTCTCTGTCATCATGATCAGACCTAT 332
 Qy 1037 GAGTTGGCAAAAGCTTT TCCAGAGGCTGAACAGGACCGGCTTCTGGCGCGCTGAAAG 1096
 Db 331 GAGTTGGCAAAAGCTTTTCCACAGGCTCAACAGGAGCAGGCTCTGGGCCACTGAAGG 272

RESULT 6

US-10-012-896-387
 ; Sequence 387, Application US/10012896
 ; Publication No. US20020183251A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stoik, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Fasih A.W.
 ; APPLICANT: Hepler, William T.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Hural, John
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: Vinals de Bassols, Carlota
 ; APPLICANT: Foy, Teresa
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Wantanabe, Yoshihiro
 ; APPLICANT: Mesgher, Madeleine Joy
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-387

Query Match 15.3%; Score 170; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 7,1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGCGCAAAAGCTTCTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGCGCAAAAGCTTCTTCCAGAGGC 120
QY 1065 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGCGCAAGGAC 1114
Db 121 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGCGCAAGGAC 170

RESULT 7

US-09-895-793-387
; Sequence 387, Application US/09895793
; Publication No. US20020192763A1

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedwick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basols, Carlota
APPLICANT: Foy, Teresa

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-387

Query Match 15.3%; Score 170; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 7,1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGCGCAAAAGCTTCTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGCGCAAAAGCTTCTTCCAGAGGC 120
QY 1065 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGCGCAAGGAC 1114
Db 121 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGCGCAAGGAC 170

RESULT 8

US-09-895-814-387
; Sequence 387, Application US/09895814
; Publication No. US20020193296A1

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedwick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-387

Query Match 15.3%; Score 170; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 7,1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 945 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGCACCAAGGACTCTTTGAGGCTTCTTCTCGATCATCAAGGCTG 60
QY 1005 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGCGCAAAAGCTTCTTCCAGAGGC 1064
Db 61 CCCCTCTCTGTCATCATGATCAGACCTATGAGTTGCGCAAAAGCTTCTTCCAGAGGC 120
QY 1065 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGCGCAAGGAC 1114
Db 121 TGAACCAAGGACCGGCTTCTGGCGGCTGAAAGGGGCAAGGCGCAAGGAC 170

RESULT 9

US-09-759-143-387
; Sequence 387, Application US/09759143

Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-387

Query Match 15.3%; Score 170; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 7.1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 945 GGGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1005 CCCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 120
QY 1065 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGCAAGGAC 1114
Db 121 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGCAAGGAC 170

RESULT 10
US-09-780-669-387
; Sequence 387, Application US/09780669
; Patent No. US2002005197A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-387

Query Match 15.3%; Score 170; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 7.1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 945 GGGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1005 CCCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 120
QY 1065 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGCAAGGAC 1114
Db 121 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGCAAGGAC 170

RESULT 11
US-09-822-827-387
; Sequence 387, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-387

Query Match 15.3%; Score 170; DB 10; Length 537;
Best Local Similarity 100.0%; Pred. No. 7.1e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 945 GGGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 1004
Db 1 GGGCCGAGTCGGGACCAAGGACTCTTTCAGGCTTCTTCCTCGGATCATCAAGGCTG 60
QY 1005 CCCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 1064
Db 61 CCCCCCTCTGTGCCATCATGATCAGACCTATGAGTTGGGAAAGCTTCTTCAGAGGC 120
QY 1065 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGCAAGGAC 1114
Db 121 TGAACCAAGACCGGCTTCTGGGGCTGAAAGGGCAAGGCAAGGAC 170

RESULT 12
US-09-908-711-46
; Sequence 46, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PA128
;; CURRENT APPLICATION NUMBER: US/09/908,711
;; CURRENT FILING DATE: 2001-07-20
;; PRIOR APPLICATION NUMBER: US01/01360
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,867
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01344
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,892
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01345
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,888
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01329
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,905
;; PRIOR FILING DATE: 2001-01-17
;; OR APPLICATION NUMBER: US01/01354
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,891
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01339
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,869
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01340
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,874
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01334
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,898
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01320
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,853
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01349
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,902
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01239
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,870
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01348
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,882
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01347
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,896
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01307
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,864
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01341
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,856
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01336
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 09/764,868
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US01/01312
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 60/179,065
;; PRIOR FILING DATE: 2000-01-31
;; PRIOR APPLICATION NUMBER: 60/180,628

;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: 60/209,467
;; PRIOR FILING DATE: 2000-06-07
;; NUMBER OF SEQ ID NOS: 167
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 46
;; LENGTH: 275
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (246)
;; OTHER INFORMATION: n equals a,t,g, or c
US-09-908-711-46

Query Match 13.8%; Score 154; DB 10; Length 275;
Best Local Similarity 89.7%; Pred. No. 9.6e-32;

Matches 174; Conservative 2; Mismatches 17; Indels 1; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACCCAGGACCCCTGCGGGCATCAGCCCTCCAGCAAAATGGT 60
|||||
DB 53 CTGAAGCTTCAAGATGGCTGACCCAGGACCCCTGCGGGCATCAGCCCTCCAGCAAAATGGT 111
|||||
QY 61 GGCCTCAGGACACCGGGGCTGTGGTTACCTCTCTTTCATGACACCCCTCGAGCTGGTGAA 120
|||||
DB 112 GGCCTCAGGACACCGGGGCTGTGGTTACCTCTCTTTCATGACACCCCTCGAGCTGGTGAA 171
|||||
QY 121 GGTTCGCTCGAGTCTTCAGCGGCCCTTCATGCGGAGGAGCTGATGCTTCTCCAGACT 180
|||||
DB 172 AGTTGCTCGAGTCTTCAGCGGCCCTTCATGCGGAGGAGCTGATGCTTCTCCAGACT 231
|||||
QY 181 GTGAGAGCTCTCCCT 194
|||||
DB 232 CTGAGAGCTCTCCCT 245
|||||

RESULT 13

US-09-960-352-2543
;; Sequence 2543, Application US/09960352
;; Patent No. US20020137139A1
;; GENERAL INFORMATION:
;; APPLICANT: Warren, Wesley C.
;; APPLICANT: Tao, Nengbing
;; APPLICANT: Byatt, John C.
;; APPLICANT: Mathialagan, Nagappan
;; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
;; FILE REFERENCE: 16511.006/37-21(10298)C
;; CURRENT APPLICATION NUMBER: US/09/960,352
;; CURRENT FILING DATE: 2001-09-24
;; NUMBER OF SEQ ID NOS: 15112
;; SEQ ID NO 2543
;; LENGTH: 246
;; TYPE: DNA
;; ORGANISM: Bos taurus
;; OTHER INFORMATION: Clone ID: 11-LIB34-070-Q1-E1-C11
US-09-960-352-2543

Query Match 10.4%; Score 116.2; DB 10; Length 246;
Best Local Similarity 87.9%; Pred. No. 9.6e-22;

Matches 138; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 1 CTGAAGCTTCAAGATGGCTGACCCAGGACCCCTGCGGGCATCAGCCCTCCAGCAAAATGGT 60
|||||
DB 65 CTGAAGCTTCAAGATGGCTGACCCAGGACCCCTGCGGGCATCAGCCCTCCAGCAAAATGGT 124
|||||
QY 61 GGCCTCAGGACACCGGGGCTGTGGTTACCTCTCTTTCATGACACCCCTCGAGCTGGTGAA 120
|||||
DB 125 GGCATCGCGCGGGGCTGTGGTTACCTCTCTTTCATGACACCCCTCGAGCTGGTGAA 184
|||||
QY 121 GGTTCGCTCGAG-TCTCAGCGGCCCTTCATGCGGAG 156
|||||
DB 185 GGTACGCTTGCAGTTCTCAGCGGCCCTCAGTGGCCAG 221
|||||

```

RESULT 14
US-09-880-107-1133/c
; Sequence 1133, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1133
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA450247
US-09-880-107-1133

Query Match      8.4%; Score 93.4; DB 10; Length 452;
Best Local Similarity 98.9%; Pred. No. 1.3e-15;
Matches 94; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1020 TCATGATCAGCACCTATGAGTTCGCGAAAGCTTCTTCAGAGGCTGAACAGGACCGGC 1079
Db 452 TCATGATCAGCACTTATGAGTTCGCGAAAGCTTCTTCAGAGGCTGAACAGGACCGGC 393

Qy 1080 TTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
Db 392 TTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 358

RESULT 15
US-09-864-761-11674/c
; Sequence 11674, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11674
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004958.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
US-09-864-761-11674

Query Match      6.7%; Score 74.8; DB 10; Length 452;
Best Local Similarity 64.4%; Pred. No. 1.1e-10;
Matches 112; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 531 TGGGCACCGTGACTGTGATCAGCCCTCGAGCTTATGCGACAAAGCTCAGGCTCAGC 590
Db 436 TTGGTCAGTAAGTGTGATAAGTCCACTAGATTGATTAGAACCAAGATGAGTCCAAGA 377

Qy 591 ATGTGTCGTACCGGAGCTGGGTGCCTGTGTTCGAACCTGCAGTGCCTCAGGCTGGCTGC 650
Db 376 AGTTTCTTACGTGGAAGTGCATCGATTGTCAGCAAGAAAGTATCTGAAGATGGTTGGA 317

Qy 651 GCTCACTGTGGCTGGGCTGGGCGCCCACTGCCCTTCGAGATGTCCTCTTCAG 704
Db 316 TTTCCCTTTGGAGGGGCTGGGCTCCTACTGTCTTTAGAGATGTACTTTCTCTCAG 263

Search completed: January 22, 2003, 19:07:54
Job time : 70.7 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2003, 19:06:17 ; Search time 1752.67 Seconds
(without alignments)
10293.906 Million cell updates/sec

Title: US-09-888-358-1
Perfect score: 1114
Sequence: 1 cgaagcttcaagatgctg.....aggcgcaaggaggcaaggac 1114

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1615406 seqs, 8097743376 residues

Tc number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estcov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Result No.	Score	Query Match	Length	DB ID	Description
1	1002.4	90.0	1402	11 AF119864	AF119864 Homo sapi
2	919.6	82.5	1020	13 BM554056	BM554056 AGENCOURT
3	848.6	76.2	1450	11 AK007934	AK007934 Mus muscu
4	831.8	74.7	1371	11 AK007276	AK007276 Mus muscu
5	826	74.1	891	9 AL522522	AL522522 AL522522
6	826	74.1	900	14 BQ883556	BQ883556 AGENCOURT

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	817.6	73.4	917	9	AL544360
8	812.6	72.9	1453	11	AK002381
9	803.8	72.2	920	14	BQ683596
10	801	71.9	1491	11	AK019396
11	798.8	71.7	981	14	BM923995
12	787	70.6	1001	9	AL519987
13	786.4	70.6	1029	13	BM459839
14	785.6	70.5	921	9	AL529906
15	780.6	70.1	1012	13	BM468665
16	777.8	69.8	1071	13	BM768227
17	777.2	69.8	969	13	BM759919
18	776.6	69.7	925	9	AL516722
19	775.4	69.6	911	9	AL528266
20	774.2	69.5	1061	14	BM923251
21	766.4	68.8	816	12	BG470240
22	764	68.6	885	9	AL514846
23	759.6	68.2	959	14	BM915493
24	755.6	67.8	939	14	BQ949458
25	755.4	67.8	965	14	BQ934841
26	753.8	67.7	862	12	BG470248
27	748.8	67.2	962	14	BQ688847
28	747.4	67.1	963	14	BQ071835
29	746.4	67.0	887	14	BQ954108
30	744.8	66.9	1474	11	AK010272
31	744.2	66.8	1077	14	BM808978
32	743.2	66.7	875	9	AL555763
33	742	66.6	1043	14	BQ073154
34	738.6	66.3	847	9	AL547590
35	737	66.2	792	12	BG470228
36	737	66.2	901	9	AL545150
37	735	66.0	976	14	BQ647782
38	734.2	65.9	961	14	BQ647042
39	733.4	65.8	1120	13	BM464148
40	731.4	65.7	1080	14	BM917004
41	729.2	65.5	918	13	BM457688
42	727.8	65.3	890	9	AL515130
43	727.2	65.3	943	13	BM018915
44	724.8	65.1	810	12	BG749617
45	722.4	64.8	994	13	BM555499

ALIGNMENTS

RESULT 1
AF119864
LOCUS AF119864 Homo sapiens PRO2163 mRNA, complete cds.
DEFINITION AF119864
ACCESSION AF119864
VERSION AF119864.1 GI:7770164
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1402)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhang, Y., Zhou, G., Bi, J., Liu, M. and He, F.
TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1402)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhang, Y., Zhou, G., Bi, J., Liu, M. and He, F.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
FEATURES
source
1. .1402
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLB8041"

```

/tissue_type="liver"
/dev_stage="fetus"
4 1017
/notes=predicted protein of H02163
/codon_start=1
/product="PRO2163"
/protein_id="AAF69618.1"
/db_xref="GI:7770165"
/translation="MVASGTVGAVVTSLEMTPLDVVKVLPQSRPMSSELSRLWS
LSYTKWKLIVGVLPYLCNGARCAWFQDPTRFTGMDAFVKI VRHSGTETLW
SGPATLVWTPATVAYTDQKAFICGREALTSDLYAPWAGALRGITVVISPL
ELMRKLOQHSVREILGACRTAVAGQWRSLMIGPPTALRDVPSALYVNEVLV
KSNMGLRPQDQTSVMSFVAGISRTVAALVTLFPDVVKTORVALGAEAVRNP
HVDSTWLLLRIRAESQTKGLFAGPLPRIIKAAPSCAIMISTVEFGKSFQRLNQDL
LGG"
BASE COUNT 261 a 432 c 401 g 308 t
ORIGIN

```

Query Match 90.0%; Score 1002.4; DB 11; Length 1402;
 Best Local Similarity 97.6%; Pred. No. 9.2e-233;
 Matches 1037; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

```

QY 53 CAATATGTCCTCAGGACCGGGCTGTGGTTACCTCTCTTCTCATGACACCCCTGGAC 112
DB 1 CAATATGTCCTCAGGACCGGGCTGTGGTTACCTCTCTTCTCATGACACCCCTGGAC 60
QY 113 GTGGTGAAGTTTGGCTGAGTCTCAGGCGCTCCATGCGCAGGAGCTGATGCTTCC 172
DB 61 GTGGTGAAGTTTGGCTGAGTCTCAGGCGCTCCATGCGCAGGAGCTGATGCTTCC 120
QY 173 TCCAGACTGTGAGGCTCTCTCTATACCAAAATGCGCTCTCTCTCAATCCACAGGGAAG 232
DB 121 TCCAGACTGTGAGGCTCTCTCTATACCAAT-----GGAAG 156
QY 233 TGCTCTCTGATGAAATGGTGTCTGAGGCTCTGTAAGTCTGCGCCAAATGGTCCCGC 292
DB 157 TGCTCTCTGATGAAATGGTGTCTGAGGCTCTGTAAGTCTGCGCCAAATGGTCCCGC 216
QY 293 TGTGCCACCTGCTTCAAGACCTTACCGCTTCACTGGCACCAGTGGTCTGCTGGAAG 352
DB 217 TGTGCCACCTGCTTCAAGACCTTACCGCTTCACTGGCACCAGTGGTCTGCTGGAAG 276
QY 353 ATGTCGAGCAGGAGGACAGGACCTCTGAGGCGGCTTCCCGCCACCTGCTGATG 412
DB 277 ATGTCGAGCAGGAGGACAGGACCTCTGAGGCGGCTTCCCGCCACCTGCTGATG 336
QY 413 ACTGTGCGAGCTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGCTGT 472
DB 337 ACTGTGCGAGCTACCGCATCTACTTCACTGCTATGACCACTGAAGGCTTCTGCTGT 396
QY 473 GGTGCGAGCCTGACCTCTGACCTTACGACACCATGGTGGCTGGCGGCTGGCGGCTG 532
DB 397 GGTGCGAGCCTGACCTCTGACCTTACGACACCATGGTGGCTGGCGGCTGGCGGCTG 456
QY 533 GGCACCGTGAAGTATGAGCCCTGGAGCTTATGCGGACAAAGCTGACGGCTCAGCAT 592
DB 457 GGCACCGTGAAGTATGAGCCCTGGAGCTTATGCGGACAAAGCTGACGGCTCAGCAT 516
QY 593 GTGTCGTACCGGAGCTGGGTGCTGTGTGCACTGAGTGGCTCAGGTCGCTGGCGC 652
DB 517 GTGTCGTACCGGAGCTGGGTGCTGTGTGCACTGAGTGGCTCAGGTCGCTGGCGC 576
QY 653 TCACTGTGCTGGGTGGGCGCCCTGACCTTCCAGATGATGCGCTTCTCAGCCCTGTAC 712
DB 577 TCACTGTGCTGGGTGGGCGCCCTGACCTTCCAGATGATGCGCTTCTCAGCCCTGTAC 636
QY 713 TGGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACAGCT 772
DB 637 TGGTTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACAGCT 696
QY 773 TCTGTGGGATGAGCTTGTGGCTGTGGCTGTCAGGAGCGGTGGCTGTCAGTGTGACT 832
DB 697 TCTGTGGGATGAGCTTGTGGCTGTGGCTGTCAGGAGCGGTGGCTGTCAGTGTGACT 756

```

```

QY 833 CTACCTTTGACGTGTAAGACCCAAAGCCAGGTGCTCTGGAGCGATGAGGCTGTG 892
DB 757 CTACCTTTGACGTGTAAGACCCAAAGCCAGGTGCTCTGGAGCGATGAGGCTGTG 816
QY 893 AGAGTGAACCCCTTGCATGTGAGCTCCACCTGGCTGCTGCGGAGGATCGGGCCGAG 952
DB 817 AGAGTGAACCCCTTGCATGTGAGCTCCACCTGGCTGCTGCTGCGGAGGATCGGGCCGAG 876
QY 953 TCGCGCACCAAGGAGCTCTTTCAGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTTCC 1012
DB 877 TCGCGCACCAAGGAGCTCTTTCAGAGGCTTCTTCTCGGATCATCAAGGCTGCCCTTCC 936
QY 1013 TGTGCCATCATGATCAGCACTTATGAGTTCCGAAAAGCTTCTTCCAGAGGCTGAACAG 1072
DB 937 TGTGCCATCATGATCAGCACTTATGAGTTCCGAAAAGCTTCTTCCAGAGGCTGAACAG 996
QY 1073 GACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1114
DB 997 GACCGGCTTCTGGCGGCTGAAAGGGGCAAGGAGGCAAGGAC 1038

```

RESULT 2
 BM554056 1020 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT.6546709 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742317
 DEFINITION 5', mRNA sequence.
 ACCESSION BM554056
 VERSION BM554056.1 GI:18793328
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1020)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLA12760 Row: e Column: 06
 High quality sequence stop: 691.
 Location/Qualifiers
 1..1020
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH MGC Library."

FEATURES
 source

BASE COUNT 175 a 327 c 302 g 213 t 3 others
 ORIGIN
 Query Match 82.5%; Score 919.6; DB 13; Length 1020;
 Best Local Similarity 98.5%; Pred. No. 1e-212;
 Matches 958; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCCCTGGGGCATCAGCCCCCTCCAGGAAATGGT 60
 Db CTGAAGCTTCAAGATGGCTGACAGGACCCCTGGGGCATCAGCCCCCTCCAGGAAATGGT 108
 QY 61 GGCCTCAGGACCGGGCTGGTGTACCTCTCTCTCATGACACCCCTGGAGCTGGTGA 120
 Db GGCCTCAGGACCGGGCTGGTGTACCTCTCTCTCATGACACCCCTGGAGCTGGTGA 168
 QY 121 GTTTCGCTCGAGTCTCAGCGCCCTCCATCGGCAGCGAGCTGATGCTTCTCCAGACT 180
 Db GTTTCGCTCGAGTCTCAGCGCCCTCCATCGGCAGCGAGCTGATGCTTCTCCAGACT 228
 QY 181 GTGAGGCTCTCTATACAAATGGCCCTCTCTCTCCATCCAGGAGAGTCCCTCT 240
 Db GTGAGGCTCTCTATACAAATGGCCCTCTCTCTCCATCCAGGAGAGTCCCTCT 288
 QY 241 GTATTGAATGGTCTCTGAGGCTCTGTACCTGTGCCCCCAATGTGTGCGCTGGCC 300
 Db GTATTGAATGGTCTCTGAGGCTCTGTACCTGTGCCCCCAATGTGTGCGCTGGCC 348
 QY 301 CTGTTTTCAAGACCCCTACCCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCGTGAG 360
 Db CTGTTTTCAAGACCCCTACCCGCTTCACTGGCACCATGGATGCTTCTGTGAAGATCGTGAG 408
 QY 361 GCACGAGGACACGAGACCCCTCTGAGGCGGCTCCCGCCACCCCTGGTGTGATGATGTC 420
 Db GCACGAGGACACGAGACCCCTCTGAGGCGGCTCCCGCCACCCCTGGTGTGATGATGTC 468
 QY 421 AGCTACCGCCATCTACTCTCACTGCTATGACCACTGAGGCTTCTCTGTGTGCTGAGC 480
 Db AGCTACCGCCATCTACTCTCACTGCTATGACCACTGAGGCTTCTCTGTGTGCTGAGC 528
 QY 481 CCTGACCTCTGACCTCTACGCAACCATGGTGGTGGCGGCTGGCGCCCTGGCCACCGT 540
 Db CCTGACCTCTGACCTCTACGCAACCATGGTGGTGGCGGCTGGCGCCCTGGCCACCGT 588
 QY 541 GACTGTGATCAGCCCTGGAGCTTATGCGGAACAGCTGCAAGCTCAGCATGTGTCGTA 600
 Db GACTGTGATCAGCCCTGGAGCTTATGCGGAACAGCTGCAAGCTCAGCATGTGTCGTA 648
 QY 601 CCGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
 Db CCGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 708
 QY 661 GCTGGCTGGGCGCCCTACCTGCGCTTCGAGATGTGCCCTTCTCAGCCCTGTACT-GGTCA 719
 Db GCTGGCTGGGCGCCCTACCTGCGCTTCGAGATGTGCCCTTCTCAGCCCTGTACT-GGTCA 768
 Q 720 ACTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGACTTCTGTGG 779
 Db ACTATGAGCTGTGAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGACTTCTGTGG 828
 QY 780 GCATGAGCTTTGTGGCTGGTGCATCTCAGGAGCGTGGCTCAGTGTGACTTACCTCT 839
 Db GCATGAGCTTTGTGGCTGGTGCATCTCAGGAGCGGCTCAGTGTGACTTACCTCT 888
 QY 840 TTGAGCTGTGTAAGACCCCAAGCCAGGTCGCTCT-GGGAGCGCATGGAGGCTGTGAGAGTG 898
 Db TTGAGCTGTGTAAGACCCCAAGCCAGGTCGCTCT-GGGAGCGCATGGAGGCTGTGAGAGTG 947
 QY 899 AACCCCTGATGTGAGCTCCACCTGGCTGCTGCGAGAGATCGGGCCGAGTCGGG 958
 Db AACCCCTGATGTGAGCTCCACCTGGCTGCTGCGAGAGATCGGGCCGAGTCGGG 1007
 QY 959 ACCAAGGAGCTCT 971
 Db ACCAAGGAGCTCT 1020
 1008 ACCAAGGAGCTCT 1020

RESULT 3
 AK007934
 LOCUS
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length

AK007934 1450 bp mRNA linear HTC 19-JAN-2002

enriched library, clone:1810062G21:homolog to HYPOTHETICAL 38.5 KDA
 PROTEIN, full insert sequence.
 AK007934
 AK007934.1 GI:12841807
 HTC; CAP trapper.
 Mus musculus (strain:CS7BL/6J) 10 day old male pancreas cDNA to
 mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
 clone:1810062G21.
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staab, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, X.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlsuki, S.
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5 (bases 1 to 1450)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
 Saito, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D.,
 Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,

KEYWORDS

SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:1700125C18.
 ORGANISM Mus musculus

REFERENCE

1 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2 Carninci,P., Shibata,K., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159

AUTHORS

Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsuoka,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipipette sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076961
 4 Kawai,J., Shingawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Akawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadoya,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M., Quackenbush,J., Schriml,L.M., Staupli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsi,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kaniya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weiss,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohsaki,S. and Hayashizaki,Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851

AUTHORS

5 (bases 1 to 1371)
 Adachi,J., Aizawa,K., Akahita,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotoani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yanunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE

JOURNAL

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

COMMENT

FEATURES

Location/Qualifiers

1..1371
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:1700125C18"
 /db_xref="MGDI:1896541"
 /db_xref="taxon:10090"
 /clone="1700125C18"
 /sex="male"
 /tissue type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 123..1133
 /notes="data source:SETR, source key:Q9UF66, evidence:ISS homolog to HYPOTHETICAL 38.5 KDA PROTEIN putative"
 /codon_start=1
 /protein_id="BAB24932.1"
 /db_xref="GI:12840722"
 /db_xref="MGDI:1915316"
 /translation="MDDQPMGTSPLQOMVAGAGAVTSLFMTPLDVVKVLSQSRP SATSELTTPSRFMSLYTKSSSALQSPGKLLYCNGLVLEPLVLPNCNCRATWQDPT RTGTLDAPVKIVHEGTRFLNSGLPVTUWVTPATATYFTAYDQKAFICGSLTSD LVAPWVAGALARMGTIVVPSFLVRLKQAOHVSRELASSVQAVTQGGKSLMUG WGTALRDPVPSALYFNELVLSWLSGLRDKDQTSVGIISFVAGISGMVAATLTLPP DVVTKQRQSLGAVEAVEVPRVDSTWLLFAESGLNLAQDLSQASSPGSSRLHPV PS"

CDS

255 a 437 c 367 g 292 t
 ORIGIN
 Query Match 74.7%; Score 831.8; DB 11; Length 1371;
 Best Local Similarity 85.6%; Pred. No. 2.5e-191;
 Matches 937; Conservative 0; Mismatches 157; Indels 1; Gaps 1;
 QY 2 TGAAGCTTCAAGATGGCTGACAGGACCTCGGGGATCATGCCCTCCAGCAAAATGGTG 61
 Db 111 TGAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 170
 QY 62 GCCTCAGGACCGGGGCTGTGGTGTACCTCTCTCTCTATGACACCTCGGACCTGGTGAAG 121
 Db 171 GCCTCAGGACCGGGGCTGTGGTGTACCTCTCTCTCTATGACACCTCGGATGATGATG 230
 QY 122 GTTCCGCTGAGTCTCAGCGGCTTCATGGCCAGCGAGCTGATGCTTCTCCAGACTG 181
 Db 231 GTCCGCTTCAGTCTCAGAGACCTCGGCAACAGCAATTCACAACTCCCTCCAGATTC 290
 QY 182 TGGAGCTCTCTTATACAAATGGCTCTCTCTCTCTCAATCCACAGGAGTGCCTCTG 241
 Db 291 TGGAGTCTCTCTTACACCAATATCTCTCCGCTCTCAGTCCCGGAGGAGTGCCTCTA 350
 QY 242 TATTGCAATGGTCTCTGGAGGCTCTGTACCTGTGCGCAAAATGGTGGCCGCTGTGCCACC 301
 Db 351 TACTGCAATGGAGTCTGGAGGCTCTGTACCTGTGCGCAATGGTGGCCGCTGTGCCACC 410
 QY 302 TGGTTTCAAGACCTTACCGCTTCTACCTGCGCACCATGATGCTTCTGTAAGATCGTAGG 361
 Db 411 TGGTTTCAAGACCTTACCGCTTCTACCTGCGCACCATGATGCTTCTGTAAGATCGTAGG 470
 QY 362 CACGAGGGCACCAGGACCTCTGAGGAGGCTCCCGCCACCTGGTGTGATGATGATGATG 421
 Db 471 CATGAGGGCCTAGGAGGCTCTGAGGAGGCTCCCGCCACCTGGTGTGATGATGATGATG 530
 QY 422 GCTACCGCCATCTTCTTCACTGCTCTATGACCAACTGAGGCTCTCTCTGTGTGAGGCC 481

```

Db 531 GCTACTGCTATCTACTTCACTGCTTACGACCACTCAAGCCCTTCTGTGGTCACTCC 590
Qy 482 CTGACCTCTGACCTCTAGCCACCATGGTGGCTGGCCGCTGGCCGCTGGGACCGTG 541
Db 591 TTGACCTCTGACCTCTAGCCACCATGGTGGCTGGCCGCTGGCCGCTGGGACCGTG 650
Qy 542 ACTGTGATCAGCCCTGAGCTTATCGGACAAAGCTCAGGCTCAGCATGTGTGTAC 601
Db 651 ACAGTTGTGAGCCCTTGGAGCTGTGGGACCAAGCTCAGGCTCAGCATGTGTGTAC 710
Qy 602 CGGAGCTGGGCTGCTGTGTTTGAAGTGTGCTGCTCAGGCTGCTGCTGCTGCTGCTG 661
Db 711 CTTGAGCTGGGCTGCTGCTGTTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770
Qy 662 CTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
Db 771 CTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830
Qy 722 TATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGCCGAGGACCAAGCTTCTGTGGG 781
Db 831 TAGAGTTGGTGAAGAGCTGGCTGAGCGAGCTGAGCAAAAGACCAAGCATCGTGGG 890
Qy 782 ATGAGCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 841
Db 891 ATCAGCTTCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 950
Qy 842 GAGTGGTAAAGACCAAGCCAGGCTGCTTGGGAGCGATGGAGGCTGTGAGAGTGAAC 901
Db 951 GATGGTGAAGACCAAGCCAGGCTGCTTGGGAGCGGCTGGAGGCTGTGAGAGTGAAG 1010
Qy 902 CCGCTGCATTTGGAGCTCCACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
Db 1011 CGGCCAGAGTGGACTCCACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1069
Qy 962 AAGGAGCTTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCTGCTGCTGCTGCTGCTG 1021
Db 1070 AGGGAGCTTTTGGAGGCTTCTTCTCGGATCATCAAGGCTGCTGCTGCTGCTGCTG 1129
Qy 1022 ATGATCAGCACTATGAGTTGGCAAAA TTCTTCAGAGGCTGAGCAAGGACCGGCTT 1081
Db 1130 ATGATCAGCACTATGAGTTGGCAAAAAGCTTCTTCCAGAGGCTCAACAGGAGCAG 1189
Qy 1082 CTGGGCGGCTGAAAG 1096
Db 1190 CTGGGCGGCTGAAAG 1204

```

```

RESULT: 5
AL522522 891 bp mRNA linear EST 13-FEB-2001
LOCUS AL522522 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF05 5
DEFINITION prime, mRNA sequence.
ACCESSION AL522522
VERSION AL522522.1 GI:12786015
KEYWORDS EST.
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 891)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/giOne="CS0DB008YF05"

```

```

FEATURES
source

```

```

/clone.lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 146 a 292 c 264 g 189 t
ORIGIN

```

```

Query Match 74.1%; Score 826; DB 9; Length 891;
Best Local Similarity 99.9%; Pred. No. 5.2e-190;
Matches 837; Conservative 0; Mismatches 0; Indels 1; Gaps :
Qy 1 CTGAAGCTTCAGATGGCTGACAGGACCTCGGGCATCAGCCCCCTCCAGAAATGCT 60
Db 55 CTGAAGCTTCAGATGGCTGACAGGACCTCGGGCATCAGCCCCCTCCAGAAATGCT 114
Qy 61 GGCCTCAGGACACCGGGGCTGTGGTTACCTCTCTCTCATGACACCCCTCGACCTGGTGA 120
Db 115 GGCCTCAGGACACCGGGGCTGTGGTTACCTCTCTCTCATGACACCCCTCGACCTGGTGA 174
Qy 121 GGTTCGCTGAGTCTCAGGGGCTCCATGCGCAGGCTGATGCTTCTCCAGACT 180
Db 175 GGTTCGCTGAGTCTCAGGGGCTCCATGCGCAGGCTGATGCTTCTCCAGACT 234
Qy 181 GTGAGCCTCTCTATACCAATTTGCCCTCTCTCCAAATCCAGAGGAGTGCCTCT 240
Db 235 GTGAGCCTCTCTATACCAATTTGCCCTCTCTCTCCAAATCCAGAGGAGTGCCTCT 294
Qy 241 GTATTGCAATGGTGTCTGAGGCTCTGTACCTGTGCCAAATGGTGCCTGTGCCAC 300
Db 295 GTATTGCAATGGTGTCTGAGGCTCTGTACCTGTGCCAAATGGTGCCTGTGCCAC 354
Qy 301 CTGTTTCAAGACCTTACCGCTTCTCTGACCATGATGATGCTTCTGTAAGATCGTGA 360
Db 355 CTGTTTCAAGACCTTACCGCTTCTCTGACCATGATGATGCTTCTGTAAGATCGTGA 414
Qy 361 GCACGAGGACACAGGACCTCTGAGAGCGGCTCTCCGCCACCTGCTGATGATGCTGCC 420
Db 415 GCACGAGGACACAGGACCTCTGAGAGCGGCTCTCCGCCACCTGCTGATGATGCTGCC 474
Qy 421 AGCTACCGCATCTACTTCTGCTGATGACCACTGAGGCTTCTGTGGTGGAGC 480
Db 475 AGCTACCGCATCTACTTCTGCTGATGACCACTGAGGCTTCTGTGGTGGAGC 534
Qy 481 CTTGACCTCTGACCTCTAGCACCCTGCTGCTGGGCGCTGGCCGCTGGGACCGT 540
Db 535 CTTGACCTCTGACCTCTAGCACCCTGCTGCTGGGCGCTGGCCGCTGGGACCGT 594
Qy 541 GACTGTGATCAGCCCTCTGAGGCTTATCGGACAAAGCTCAGGCTCAGCATGTCTGTA 600
Db 595 GACTGTGATCAGCCCTCTGAGGCTTATCGGACAAAGCTCAGGCTCAGCATGTCTGTA 654
Qy 601 CCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 655 CCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
Qy 661 GCTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 715 GCTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 773
Qy 721 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCTGAGGACCAAGCTTCTGTGG 780
Db 774 CTATGAGCTGGTGAAGAGCTGGCTCAATGGGCTCAGGCTGAGGACCAAGCTTCTGTGG 833

```

```

QY 781 CATGAGCTTTGGCTGTGCTGCTCAGGAGCGGTGGCTCAGTGTGACTTACCC 838
Db 834 CATGAGCTTTGGCTGTGCTGCTCAGGAGCGGTGGCTCAGTGTGACTTACCC 891

RESULT 6
BQ883556
LOCUS DEFINITION BQ883556 900 bp mRNA linear EST 16-AUG-2002
AGENCOURT 8727656 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6340846
5', mRNA sequence.
ACCESSION BQ883556
VERSION BQ883556.1 GI:22275564
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 900)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: ccgaps-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2537 row: o column: 23
High quality sequence stop: 629.
1. .900
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6340846"
/clone_lib="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
1 COUNT 150 a 286 c 262 g 200 t
ORIGIN

Query Match 74.1%; Score 826; DB 14; Length 900;
Best Local Similarity 97.1%; Pred. No. 5.2e-190;
Matches 873; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 116 GTGAGGTTTCGCTGAGTCTCAGCGGCCCTCCATGCGCAGAGTGATGCTTCCTCC 175
Db 1 GTGAGGTTTCGCTGAGTCTCAGCGGCCCTCCATGCGCAGAGTGATGCTTCCTCC 60

QY 176 AGACTGTGAGGCTCTCCCTATACAAATGCGCTCTCTCCATCCACAGGAAGTGC 235
Db 61 AGACTGTGAGGCTCTCCCTATACAAATGCGCTCTCTCCATCCACAGGAAGTGC 120

QY 236 CTCCTGTATTGCAATGGTGTCTGGAGGCTCTGTACCTGTGCGCAAAATGGTGGCGGTGT 295
Db 121 CTCCTGTATTGCAATGGTGTCTGGAGGCTCTGTACCTGTGCGCAAAATGGTGGCGGTGT 180

QY 296 GCGACCTGGTTTCAAGACCTTACCGCTTCTACTGGACCATGATGCTTCTGTGAAGATC 355
Db 181 GCGACCTGGTTTCAAGACCTTACCGCTTCTACTGGACCATGATGCTTCTGTGAAGATC 240

QY 356 GTGAGGCAACGAGGGCACCAGGACCTCTGGAGCGGCTCCCGCCACCCCTGGTGTGACT 415

```

```

Db 241 GTGAGGCAACGAGGGCACCAGGACCTCTGGAGCGGCTCCCGCCACCCCTGGTGTGACT 300
QY 416 GTGCGAGCTTACCGCATCTACTTCACTGCGCTATGACCAACTGAAGGCTTCTCTGTGTGT 475
Db 301 GTGCGAGCTTACCGCATCTACTTCACTGCGCTATGACCAACTGAAGGCTTCTCTGTGTGT 360
QY 476 CGAGCCCTGACCTCTGACCTCTAGCGACCATGCTGGGCTGGGCGGCTGGCCCGCTGGGC 535
Db 361 CGAGCCCTGACCTCTGACCTCTAGCGACCATGCTGGGCTGGGCGGCTGGCCCGCTGGGC 420
QY 536 ACCGTGACTGTGATCAGCCCTCTGGAGCTTATCGCGACAAAGCTGCAGGCTCAGCATGTG 595
Db 421 ACCGTGACTGTGATCAGCCCTCTGGAGCTTATCGCGACAAAGCTGCAGGCTCAGCATGTG 480
QY 596 TCGTACCGGAGCTGGTGGTCTGTGTGAACTGCGAGTGGTCAAGGCTGGCTGGCTCA 655
Db 481 TCGTACCGGAGCTGGTGGTCTGTGTGAACTGCGAGTGGTCAAGGCTGGCTGGCTCA 540
QY 656 CTGTGCTGGGCTGGGCGGCCACTGCGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGG 715
Db 541 CTGTGCTGGGCTGGGCGGCCACTGCGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGG 600
QY 716 TTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGCTCAGGCCGAAAGGACCAAGCTTCT 775
Db 601 TTCAACTATGAGCTGGTGAAGAGCTGGCTCAATGGCTCAGGCCGAAAGGACCAAGCTTCT 660
QY 776 GTGGGATGAGCTTTGTGGTGGTGGCATCTCAGGACGGTGGCTGCAGTGTGACTCTA 835
Db 661 GTGGGATGAGCTTTGTGGTGGTGGCATCTCANGACGGTGGCTGCAGTGTGACTCTA 720
QY 836 CCTTTGACGTGTGAAAGACCCAGCCAGCTGCTGGAGCGATGGAGGCTGTGAGA 895
Db 721 CCTTTGACGTGTGAAAGACCCAGCCAGCTGCTGGAGCGATGGAGGCTGTGAGA 780
QY 896 GTGAACCCCTCGATGTGAGCTCCACCTGG-CTGCTGCTGGG-AGGATCCGGGCGAGT 953
Db 781 GTGAACCCCTCGATGTGAGCTCCACCTGGCTGCTGCTGGGAGAAACCCGGGCGAGT 840
QY 954 CGGGCACCAAGGGA---CTTTTGCAGGCTCTCTTCCTCGGATCATCAAGGCTGCCCC 1009
Db 841 CCGGCACCAAGGGAACCTTTTTCAGGCTCTCTTCCTCGGATCATCAAGGCTGCCCC 899

RESULT 7
AL544360 917 bp mRNA linear EST 16-FEB-2001
LOCUS AL544360 LTT_NFL006.PL2 Homo sapiens cDNA clone CSODI018YN07 5
prime, mRNA sequence.
ACCESSION AL544360
VERSION AL544360.1 GI:12876840
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 917)
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI018YN07"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT)-primer. Five prime end

```

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 153 a 293 c 268 g 195 t 8 others
ORIGIN

Query Match 73.4% Score 817.6; DB 9; Length 917;
Best Local Similarity 98.5%; Pred. No. 5.8e-188;
Matches 847; Conservative 8; Mismatches 2; Indels 3; Gaps 3;

QY 1 CTGAAGCTTCAGATGGCTGACACGAGACCTCGGGGCAATGAGCCCTCCAGCAATGGT 60
DB 55 CTGAAGCTTCAGATGGCTGACACGAGACCTCGGGGCAATGAGCCCTCCAGCAATGGT 114
QY 61 GGCTTCAGGACCGGGGCTGTGGTTACCTCTCTTCATGACACCCCTGACGCTGTGA 120
DB 15 GGCTTCAGGACCGGGGCTGTGGTTACCTCTCTTCATGACACCCCTGACGCTGTGA 174
QY 121 GGTTCGCTGAGTCTAGCGGGGCTCCATGCGCAGGAGCTGATGCTTCTCCAGACT 180
DB 175 GGTTCGCTGAGTCTAGCGGGGCTCCATGCGCAGGAGCTGATGCTTCTCCAGACT 234
QY 181 GTGGAGGCTCTCTATACCAATGGCTCTCTCTCTCAATCCAGGGAGTGGCTCTCT 240
DB 235 GTGGAGGCTCTCTATACCAATGGCTCTCTCTCTCAATCCAGGGAGTGGCTCTCT 294
QY 241 GTATTGCAATGGTCTCTGAGGCTCTGATGCTGCGCAATGGTGGCGCTGTCAC 300
DB 295 GTATTGCAATGGTCTCTGAGGCTCTGATGCTGCGCAATGGTGGCGCTGTCAC 354
QY 301 CTGGTTTCAAGACCTTACCGCTTCTACTGGACCATGATGATGCTTCTGTAAGATCGT 360
DB 355 CTGGTTTCAAGACCTTACCGCTTCTACTGGACCATGATGATGCTTCTGTAAGATCGT 414
QY 361 GCACAGGCGCAGGAGGCTCTGAGGAGGCTCTCCGCGCAGGCTGATGATGCTGCT 420
DB 415 GCACAGGCGCAGGAGGCTCTGAGGAGGCTCTCCGCGCAGGCTGATGATGCTGCT 474
QY 421 AG-CTACCGCATCTACTTCACTGCTCTATGACCAACTGAAGGCTCTCTGCTGCTG 479
DB 475 AGACTACCGCATCTACTTCACTGCTCTATGACCAACTGAAGGCTCTCTGCTGCTG 534
QY 480 CCCTGACCTCTGACCTTACGCAACCATGGTGGTGGCGGCTGGCGGCTGGGACCG 539
DB 535 CCCTGACCTCTGACCTTACGCAACCATGGTGGTGGCGGCTGGCGGCTGGGACCG 594
QY 540 TGACTGTGATCAGCCCTCGAGCTTATCGGCAAAAGCTGAGGCTCAGATGCTGCT 599
DB 595 TGACTGTGATCAGCCCTCGAGCTTATCGGCAAAAGCTGAGGCTCAGATGCTGCT 654
QY 600 ACCGGAGCTGGGTGGCTGTGTTTCAACTGAGTGGCTCAGGCTGCTGCGCTCACTGT 659
DB 655 ACCGGAGCTGGGTGGCTGTGTTTCAACTGAGTGGCTCAGGCTGCTGCGCTCACTGT 714
QY 660 GGCTGGGCTGGGCGCCACATGCGCTTCAGATGTCGCTTCTCAGCCCT-CTACTGGTTC 718
DB 715 GGCTGGGCTGGGCGCCACATGCGCTTCAGATGTCGCTTCTCAGCCCTGGTACTGGTTC 774
QY 719 AACTATGAGCTGTGAGAGCTGGCTCAATGGCTCAGGCGGAGGACAGACTTCTGTG 778
DB 775 AACTATGAGCTGTGAGAGCTGGCTCAATGGCTCAGGCGGAGGACAGACTTCTGTG 833
QY 779 GGCATGAGCTTTGTGGCTGGTGGCTCTCAGGAGCGTGGCTCAGTGGTCACTTACCC 838
DB 834 GGCATGAGCTTTGTGGCTGGTGGCTCTCAGGAGCGTGGCTCAGTGGTCACTTACCC 893
QY 839 TTTGAGCTGTAAGACCCA 858
|||||

Db 894 TTTGACCTGCTAAGACCA 913

RESULT 8

AK002381

LOCUS

DEFINITION

AK002381

1453 bp mRNA linear

Mus musculus adult male kidney cDNA, RIKEN full-length enriched

library, clone:0610009E11:homolog to HYPOTHETICAL 38.5 KDA PROTEIN,

full insert sequence.

AK002381

AK002381.1 GI:12832321

HTC; CAP trapper.

Mus musculus (strain:C57BL/6J) adult male kidney cDNA to mRNA,

clone.lib:RIKEN full-length enriched mouse cDNA library

clone:0610009E11.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Ogawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076961

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, F., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,

Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Badarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,

Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,

Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,

and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5 (bases 1 to 1453)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,

Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,

Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, K., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

Direct Submission

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGGAGAGAGCGCGCGAATCGAGTGTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Ret = 5.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGGAGAGAGAGTCCAGAGCTCAATTAATTAACCCCGCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

COMMENTS

Location/Qualifiers

1. .1453

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:0610009E11"

/db_xref="MGD:MGI:1892017"

/db_xref="taxon:10090"

/clone="0610009E11"

/sex="male"

/tissue type="kidney"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

1. .1453

/note="data source:SPTR, source Key:Q9UF66, evidence:ISS homolog to HYPOTHETICAL 38.5 KDA PROTEIN"

/db_xref="MGD:MGI:1915316"

FEATURES

source

Query Match 72.9%; Score 812.6; DB 11; Length 1453;
Best Local Similarity 85.7%; Pred. No. 1.2e-186;
Matches 938; Conservative 0; Mismatches 154; Indels 3; Gaps 3;

BASE COUNT 268 a 456 c 403 g 326 t

ORIGIN

2 TGAAGCTTCAAGATGGCTGACAGGACCCCTGGGGATCAGCCCCCTCCAGCAATGGTG 61
Db |
90 TGAATTCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 149
Qy 62 GCCTCAGGACCGGGCTGTGGTACCTCTCTCTTCATGACACCCCTGGACGTGGTGAAG 121
Db 150 GCCTCAGGACCGGGCTGTGGTACCTCTCTCTTCATGACACCCCTGGACGTGGTGAAG 209
Qy 122 GTTCGCTCAGTCTCAGCGGCCCTCCATGCGCAGCGAGCTGATGCTCTCTCCACACTG 181
Db 210 GTCCGCTCAGTCTCAGCGGCCCTCGGCAACGCGAATGACAACTCCCTCCAGATTC 269
Qy 182 TGGAGCTCTCTATACCAATGCGCTCTCTCTCCAAATCCACAGGGAAGTCCCTCTG 241
Db 270 TGGAGCTCTCTATACCAATGCGCTCTCTCTCCAAATCCACAGGGAAGTCCCTCTA 329

Qy 242 TATGCAATGGTGTCTGGAGCCCTGTACCTGTGCCCCAAATGGTCCCGCTGTGCACC 301
Db |
330 TACTGAATGAGTCTGGAGCCCTGTACCTGTGCCCCAAATGGTACCCTGTGCACC 389
Qy 302 TGGTTTCAAGACCTTACCCGCTTCACTGGCAACCAATGATGCTTCTGTGAAGATCTGAGG 361
Db 390 TGGTTTCAAGACCTTACCCGCTTCACTGGCAACCTTGTGAGATCTGTGCGG 449
Qy 362 CACGAGGACACAGGACCTTGGAGCGCCCTCCCGCACCTCTGTGATGATGATGATG 421
Db 450 CATGAGGACATAGGACCTTGGAGCGCCCTCCCGCACCTCTGTGATGATGATGATG 509
Qy 422 GCTACCGGACCTTACTTCACTGCTTATGACCAACTGGAAGCCCTTCTGTGTGGTGCAGC 481
Db 510 GCTACTGTATCTTCACTGCTTACGACCAACTGGAAGCCCTTCTGTGTGGTGCAGTCC 569
Qy 482 CTGACCTCTGACCTTACGCAACCAATGATGCTGCGCGCTGGCGCCCTGGGCAACCTGTG 541
Db 570 TTGACCTCTGACCTTACGCAACCAATGATGCTGCGCGCTGGCGCCCTGGGCAACCTGTG 628
Qy 542 ACTGTGATCAGCCCGCTTGGAGCTTATGCGGCAAAAGCTGAGGCTCAGCATGTGCTGATC 601
Db 629 ACAGTGTGAGCCCGCTTGGAGCTTGTGCGGCAAAAGCTGAGGCTCAGCATGTGCTGATC 688
Qy 602 CCGGAGCTGGGTGCTGTGTTGAACTGCACTGCTGCTCAGGCTGGGCTGGGCTGCTGCTG 661
Db 689 CCGGAGCTGGGTGCTGTGTTGAACTGCACTGCTGCTCAGGCTGGGCTGGGCTGCTGCTG 748
Qy 662 CTGGGCTGGGCGCCCACTGCTTCTCGAGATGCTGCTTCTCAGCCCTGTGCTGCTGCTG 721
Db 749 CTGGGCTGGGCGCCCACTGCTTCTCGAGATGCTGCTTCTCAGCTGTGCTGCTGCTGCT 807
Qy 722 TATGAGCTGTGGAAGAGCTGGCTCAATGGGCTCAGGCGGCAAGGACAGACTTCTGTGGGC 781
Db 808 TACAGTGTGGAAGAGCTGGCTCAGGCGGCAAGGACAGACTTCTGTGGGC 867
Qy 782 ATGAGCTTTTGGCTGGTGGCACTCTCAGGAGCGGTGGCTGCGATGCTGCTGCTGCTGCT 841
Db 868 ATGAGCTTTTGGCTGGTGGCACTCTCAGGAGCGGTGGCTGCGATGCTGCTGCTGCTGCT 927
Qy 842 GACGTGTAAGACCAACCGCAGGTGCTCTGCGAGCGGATGAGGCTGTGAGAGTGAAC 901
Db 928 GATGTGTGAAGACCAACCGCAGGTGCTCTGCGAGCGGTGAGGCTGTGAGAGTGAAG 987
Qy 902 CCCCTGATCTGAGCTTCCACCTGGCTGCTGCGGAGGATCCGGGCGGAGTCCGGGCAAC 961
Db 988 CCGCCAGAGTGGACTCCACCTGGCTTCTG-TTCGCAAGTCCGGGCTGAATCTGGCAAC 1046
Qy 962 AAGGAGCTCTTTGCGAGGCTTCTCTCTCGGATCATCAAGGCTGCCCCCTCTGTCGCTC 1021
Db 1047 AGGAGCTCTTTGCGAGGCTTCTCTCTCGGATCATCAAGGCTGCCCCCTCTGTCGCTC 1106
Qy 1022 ATGATCAGCACTTATGATTCGGCAAAAGCTTCTTCCAGAGGCTGAACAGGACCGGCTT 1081
Db 1107 ATGATCAGCACTTATGATTCGGCAAAAGCTTCTTCCAGAGGCTCAACAGGAGCAGCT 1166
Qy 1082 CTGGGCGGTGAAG 1096
Db 1167 CTGGGCGGTGAAG 1181

RESULT 9

B0683596

LOCUS

DEFINITION

AGENCOURT 8059897 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091198

5', mRNA sequence.

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

B0683596

AGENCOURT 8059897 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091198

5', mRNA sequence.

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

920 bp

mrna

linear

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

920 bp

mrna

linear

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

920 bp

mrna

linear

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

920 bp

mrna

linear

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

EST 15-JUL-2002

REFERENCE
 AUTHORS NIH-MGC http://imgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapps-remail.nih.gov
 Tissue Procurement: DCTD/STP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLCM2333 row: m column: 23
 High quality sequence stop: 697.
 Location/Qualifiers
 1..920
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6091198"
 /clone_lib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dr priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley), using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 161 a 288 c 273 g 196 t 2 others
 ORIGIN

Query Match 72.2%; Score 803.8; DB 14; Length 920;
 Best Local Similarity 96.8%; Pred. NO. 1.3e-184;
 Matches 850; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

QY 1 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGCT 50
 DB 31 CTGAAGCTTCAAGATGGCTGACAGGACCTCGGGCATCAGCCCTCCAGCAATGCT 90
 QY 61 GGCTCAGGACCGGGCTGTGGTACTCTCTTCATGACACCCCTGGACGTGTGAA 120
 DB 91 GGCTCAGGACCGGGCTGTGGTACTCTCTTCATGACACCCCTGGACGTGTGAA 150
 QY 121 GGTTCGCTGAGCTCTAGCGGCGCTCCATGGCCAGGAGTGTCTTCTCCAGACT 180
 DB 151 GGTTCGCTGAGCTCTAGCGGCGCTCCATGGCCAGGAGTGTCTTCTCCAGACT 210
 QY 181 GTGAGGCTCTCTATACCAATTTGCCCTCTCTCTCAATCC;CAGGGAAGTGCCTCT 240
 DB 211 GTGAGGCTCTCTATACCAAT-----GGAGTGCCTCT 246
 QY 241 GTATTGCAATGGTCTGAGGCTCTGTACTGTGCCAATGGTGCCTGTGSCAC 300
 DB 247 GTATTGCAATGGTCTGAGGCTCTGTACTGTGCCAATGGTGCCTGTGSCAC 306
 QY 301 CTGGTTTCAAGACCTACCCGCTTCTACTGSCACCATGATGCTTGTGAGATGTGAG 360
 DB 307 CTGGTTTCAAGACCTACCCGCTTCTACTGSCACCATGATGATGCTTGTGAGATGTGAG 366
 QY 361 GCACGAGGCGACAGGACCTCTGGAGCGGCTCCGCCACCTGGTGTGATGACTGTGCC 420
 DB 367 GCACGAGGCGACAGGACCTCTGGAGCGGCTCCGCCACCTGGTGTGATGACTGTGCC 426
 QY 421 AGCTACCGGATCTACTTCTAGCTGATGACCACTGAAGGCTTCTGTGTGTCGAGC 480
 DB 427 AGCTACCGGATCTACTTCTAGCTGATGACCACTGAAGGCTTCTGTGTGTCGAGC 486
 QY 481 CCTGACCTCTGACCTCTACGCCACCATGGTGGCTGGCGCTGGCGCTGGGACCGT 540
 DB 487 CCTGACCTCTGACCTCTACGCCACCATGGTGGCTGGCGCTGGCGCTGGGACCGT 546

QY 541 GACTGTGATCAGCCCTCGAGGCTTATCGGACAAAGCTGAGGCTCAGATGTGTCTA 600
 DB 547 GACTGTGATCAGCCCTCGAGGCTTATCGGACAAAGCTCAGGCTCAGATGTGTCTA 606
 QY 601 CCGGAGCTGGGTGCTGTGTTCGAACCTGCACTGCGCTCAGGCTGGCTGGCGCTCACTGTG 660
 DB 607 CCGGAGCTGGGTGCTGTGTTCGAACCTGCACTGCGCTCAGGCTGGCTGGCGCTCACTGTG 666
 QY 661 GCTGGGCTGGGCGCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 720
 DB 667 GCTGGGCTGGGCGCCCACTGCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGTTCAA 726
 QY 721 CTATGAGCTGTGAGAGGCTGCTCAATGGGCTCAGCGCAGAGGACCACTTCTGTGGG 780
 DB 727 CTATGAGCTGTGAGAGGCTGCTCAATGGGCTCAGCGCAGAGGACCACTTCTGTGGG 786
 QY 781 CATGAGCTTGTGGCTGGTGGCATCTCAGGACGCTGGCTGCACTGCTTACCTACCC-T 839
 DB 787 CATGAGCTTGTGGCTGGTGGCATCTCAGGACGCTGGCTGCACTGCTTACCTACCC-T 846
 QY 840 TTGACGTGTAAAGACCCCAACCGAGCTCGCTCTGGGA 877
 DB 847 TTGACGTGTAAAGACCCCAACCGAGCTCGCTCTGGGA 884

RESULT 10
 AK019396
 LOCUS
 DEFINITION Mus musculus 12 days embryo head cDNA, RIKEN full-length enriched library, clone:3010027G13:homolog to HYPOTHETICAL 38.5 KDA PROTEIN, full insert sequence.
 ACCESSION AK019396
 VERSION AK019396.1 GI:12859581
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:O57BL/6J) 12 days embryo head cDNA to mRNA, cDNA library
 clone:3010027G13.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Iizawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Iizawa, M., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamahata, I., Saito, T., Okazaki, Y., Gojobori, I., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fieschmann, W., Gaasterland, T., Gissi, C., King, B., Kocchiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Buit, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, C., Rodriguez, I., Sakamoto, N., Sakaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5 (bases 1 to 1491)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Buit, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hitamoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawaji, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Salto, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sakaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGAGTCTAATTAATCCCTCCCTCCCTCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

Location/Qualifiers

1. .1491
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:3010027G13"
/db_xref="MGD:MGI:1899102"
/db_xref="taxon:10090"
/clone="3010027G13"
/tissue type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"

misc_feature

1. .1491
/note="data source:SPTR, source key:Q9UF66, evidence:ISS homolog to HYPOTHETICAL 38.5 KDA PROTEIN"
/db_xref="MGD:MGI:1915316"

TITLE
JOURNAL

COMMENT

FEATURES
source

ORIGIN	270 a	470 c	418 g	330 t	3 others
Query Match	71.9%	Score 801;	DB 11;	Length 1491;	
Best Local Similarity	85.5%	Pred. No. 8e-184;			
Matches 938;	Conservative	0;	Mismatches 155;	Indels	4; Gaps
QY	2	TGAAGCTTCAAGATGCTCACCAGAGACCTCGCGGACATCAGCCCTCCAGCAATAGGTG	61		
Db	184	TOAATTCGAAGATGATGATCAGACCTCGGGGATAGTCCCTCCAGCAATAGGTG	183		
QY	62	GCCTCAGCACCGGGGCTCTGGTTACCTCTCTCATGACACCTCGAGCTGGTGAAG	121		
Db	184	GCCTCAGACCGGGGCTCTGGTTACCTCTCTCATGACACCTCGAGCTGGTGAAG	243		
QY	122	GTTGCTTCAGTCTCAGCGCCCTCCATCGCCAGGAGCTGATGCCTTCCTCCAGACTG	181		
Db	244	GTCCGCTTCAGTCTCAGAGACCTCGGCAACAGCGAATTCGAACTCCCTCCAGATTC	303		
QY	182	TGGAGCTCTCTCTATACCAATATGCTCTCTCTCAATCCAGAGGAATGCTCTCTG	241		
Db	304	TGGAGCTCTCTCTATACCAATATGCTCTCTCTCAATCCAGAGGAATGCTCTCTG	363		
QY	242	TATTGCAATGGTGTCTCTGGAGCTCTCTACCTGTGCCAAATGGTGCCTGTGCCACC	301		
Db	364	TACTGCAATGGAGTCTCTGGAGCTCTCTACCTGTGCCAAATGGTGCCTGTGCCACC	423		
QY	302	TGTTTCAAGACCTTACCCGCTTCACTGGCACTTGGTGCCTCTGTAAGATCGTGAGG	361		
Db	424	TGTTTCAAGACCTTACCCGCTTCACTGGCACTTGGTGCCTCTGTAAGATCGTGAGG	483		
QY	362	CACGAGGCAACAGAGACCTCTGGAGCGGCTCTCCCGACCTCCCTGGTGTGATGTCGA	421		
Db	484	CATGAGGCACTAGAGCCTCTGGAGCGGCTCTCCCGACCTCCCTGGTGTGATGTCGA	543		
QY	422	GCTACCGCACTTACTTCACTGCTATGACCAACTGAAGGCTCTCTGTGTGTCGAGCC	481		
Db	544	GCTACTGCTATCTACTTCACTGCTATGACCAACTGAAGGCTCTCTGTGTGTCGAGCC	603		
QY	482	CTGACCTCTGACCTCTAGCACCTCTGGTGGTGGCGCTGGGCGCTGGGCGCTGGGCG	541		
Db	604	TTGACCTCTGACCTCTAGCACCTCTGGTGGTGGCGCTGGGCGCTGGGCGCTGGGCG	663		
QY	542	ACTGTGATCAGCCCTCTGGAGCTTATGCGGACAAGCTGAGGCTCAGCATGTGTCTAC	601		
Db	664	ACATGTGTCAGCCCTCTGGAGCTTATGCGGACAAGCTGAGGCTCAGCATGTGTCTAC	723		
QY	602	CGGAGCTGGTGGCTGTGTTCGAACTGCACTGAGTGGCTCAGGCTGGCTGGGCTCAGCT	661		
Db	724	CGTGAAGT-GCTTCTCTGTTCGAACTGCACTGAGTGGCTCAGGCTGGCTGGGCTCTCT	782		
QY	662	CTGGCTGGGCGCCCATG-CCTTCGAGATGTGCTTCTCAGCCCTGTACTGTGTTCAA	720		
Db	783	CTGGCTGGGCGCCCATG-CCTTCGAGATGTGCTTCTCAGCCCTGTACTGTGTTCAA	842		
QY	721	CTATGAGCTGTGAAGAGCTGCTCAATGGGCTCAGGCGCAAGACACAGCTTCTGTGGG	780		
Db	843	CTACAGTGTGTGAAGAGCTGCTCAATGGGCTCAGGCGCAAGACACAGCTTCTGTGGG	902		
QY	781	CATGAGCTTTGTGGCTGGTGGCATCTCAG-GGACGGTGGCTGAGTGTGCTTACCTCT	839		
Db	903	CATGAGCTTTGTGGCTGGTGGCATCTCAG-GGACGGTGGCTGAGTGTGCTTACCTCT	962		
QY	840	TTGAGCTGGTAAAGACCCCAACCCAGTGTCTCTGGAGCGATGGAGGCTGTGAGAGTGA	899		
Db	963	TCGATGTGGTAAAGACCCCAACCCAGTGTCTCTGGAGCGATGGAGGCTGTGAGAGTGA	1022		
QY	900	ACCCCTGCATGTGGAGCTCCAGTGGCTGTCTGTGGAGGATTCGGGCGAGTTCGGCA	959		
Db	1023	AGCCGCGCAGTGTGGAGCTCCAGTGGCTGTCTGTGGAGGATTCGGGCGAGTTCGGCA	1081		
QY	960	CAAAGGAGCTCTTTTGGAGGCTTCTTCTCGGATCATCAAGCTGCCCTCTCTGTCGA	1019		

Search completed: January 23, 2003, 00:15:26
Job time : 1772.67 secs